

```

1 ATGGATTTTC TCCTGGCGCT GGTGCTGGTA TCCTCGCTCT ACCTGCAGGC
51 GGCCGCCGAG TTCGACGGGA GGTGGCCAG GCAAATAGTG TCATCGATTG
101 GCCTATGTCG TTATGGTGGG AGGATTGACT GCTGCTGGGG CTGGGCTCGC
151 CAGTCTTGGG GACAGTGTC A GCTTTCTAC GTCTTAAGGC AGAGAATAGC
201 CAGGATAAAG TGCCAGTGTA AAGCTGTGTG CCAACCACGA TGCAAACATG
251 GTGAATGTAT CGGGCCAAAC AAGTGCAAGT GTCATCCTGG TTATGCTGGA
301 AAAACCTGTA ATCAAGACGA GCACATCCCA GTCCTCTTG ACCAAGGCAG
351 TGAACAGCCT CTTTTCCAAC CCCTGGATCA CCAAGCCACA AGTTTGCCTT
401 CAAGGGATCT AAATGAGTGT GGCTGAAGC CCCGGCCCTG TAAGCACAGG
451 TGCATGAACA CTTACGGCAG CTACAAGTGC TACTGTCTCA ACGGATATAT
501 GCTCATGCCG GATGGTTCCT GCTCAAGTGC CCTGACCTGC TCCATGGCAA
551 ACTGTCAAGT TGGCTGTGAT GTTGTAAAG GACAAATACG GTGCCAGTGC
601 CCATCCCCTG GCCTGCAGCT GGCTCCTGAT GGGAGGACCT GTGTAGATGT
651 TGATGAATGT GCTACAGGAA GAGCCTCCTG CCCTAGATTT AGGCAATGTG
701 TCAACACTTT TGGGAGCTAC ATCTGCAAGT GTCATAAAGC CTTGATCTC
751 ATGTATATTG GAGGCAAATA TCAATGTCAT GACATAGACG AATGCTCACT
801 TGGTCAGTAT CAGTGCAGCA GCTTTGCTCG ATGTTATAAC GTACGTGGGT
851 CCTACAAGTG CAAATGTAAA GAAGGATACC AGGGTGATGG ACTGACTTGT
901 GTGTATATCC CAAAAGTTAT GATTGAACCT TCAGGTCCAA TTCATGTACC
951 AAAGGGAAAT GGTACCATTT TAAAGGGTGA CACAGGAAAT AATAATTGGA
1001 TTCCTGATGT TGGAAGTACT TGGTGGCCTC CGAAGACACC ATATATTCCT
1051 CCTATCATTA CCAACAGGCC TACTTCTAAG CCAACAACAA GACCTACACC
1101 AAAGCCAACA CCAATTCCTA CTCCACCACC ACCACCACC CTGCCAACAG
1151 AGCTCAGAAC ACCTTACCA CCTACAACCC CAGAAAGGCC AACCACCGGA
1201 CTGACAACTA TAGCACCAGC TGCCAGTACA CCTCCAGGAG GGATTACAGT
1251 TGACAACAGG GTACAGACAG ACCCTCAGAA ACCCAGAGGA GATGTGTTC
1301 TTCCACGGCA ACCTTCAAAT GACTTGTTTG AAATATTTGA AATAGAAAGA
1351 GGAGTCAGTG CAGACGATGA AGCAAAGGAT GATCCAGGTG TTCTGGTACA
1401 CAGTTGTAAT TTTGACCATG GACTTTGTGG ATGGATCAGG GAGAAAGACA
1451 ATGACTTGCA CTGGGAACCA ATCAGGGACC CAGCAGGTGG ACAATATCTG
1501 ACAGTGTGCG CAGCCAAAGC CCCAGGGGGA AAAGCTGCAC GCTTGGTGCT
1551 ACCTCTCGGC CGCCTCATGC ATTACAGGGA CCTGTGCCTG TCATTACGGC
1601 ACAAGGTGAC GGGGCTGCAC TCTGGCACAC TCCAGGTGTT TGTGAGAAAA
1651 CACGGTGCCC ACGGAGCAGC CCTGTGGGGA AGAAATGGTG GCCATGGCTG
1701 GAGGCAAACA CAGATCACCT TGCGAGGGGC TGACATCAAG AGCGTCGTCT
1751 TCAAAGGTGA AAAAAGGCGT GGTCACTG GGGAGATTGG ATTAGATGAT
1801 GTGAGCTTGA AAAAAGGCCA CTGCTCTGAA GAACGCTAA (SEQ ID NO:1)

```

FEATURES:

Start Codon: 1
Stop Codon: 1837

Homologous proteins:

Top 10 BLAST Hits:

	Score	E
CRA 50000098943299 /altid=gi 15795193 /def=ref NP_277060.1 nep...	1190	0.0
CRA 46000103872918 /altid=gi 15128103 /def=gb AAK84391.1 AF3970...	1147	0.0
CRA 84000042916333 /altid=gi 15430248 /def=gb AAK96011.1 (AY03...	1128	0.0
CRA 84000042916331 /altid=gi 15430246 /def=gb AAK96010.1 (AY03...	1084	0.0
CRA 78000201656899 /altid=gi 14763663 /def=ref XP_040394.1 epi...	397	e-109
CRA 163000000492958 /altid=gi 9506563 /def=ref NP_062270.1 EGF...	396	e-109
CRA 157000140328010 /altid=gi 6752658 /def=gb AAF27812.1 AF1860...	396	e-109
CRA 157000140333453 /altid=gi 13124888 /def=ref NP_056322.2 ep...	395	e-109
CRA 1000682328934 /altid=gi 7512796 /def=pir T17324 hypothetical...	385	e-106
CRA 60000046724650 /altid=gi 13929180 /def=ref NP_114014.1 fib...	147	4e-34

FIGURE 1A

Blast hits to dbEST:

<u>CRA Number</u>	<u>gi Number</u>	<u>Score</u>	<u>Expect</u>
CRA 113000119195127	gi 14391046	1068 bits (539)	0.0
CRA 110000024586745	gi 12320908	866 bits (437)	0.0
CRA 87000000737360	gi 7312275	842 bits (425)	0.0
CRA 113000119269070	gi 14399233	831 bits (419)	0.0
CRA 113000119283425	gi 14400823	827 bits (417)	0.0
CRA 113000119283416	gi 14400822	785 bits (396)	0.0
CRA 112000057002060	gi 14377756	680 bits (343)	0.0
CRA 147000029699826	gi 11300389	668 bits (337)	0.0
CRA 117000088322182	gi 12462132	563 bits (284)	1e-157

Expression (Tissue Source):

<u>gi Number</u>	<u>Organ</u>
gi 14391046	head_neck
gi 12320908	nervous_tumor
gi 7312275	head_neck
gi 14399233	head_neck
gi 14400823	head_neck
gi 14400822	head_neck
gi 14377756	colon
gi 11300389	breast_normal
gi 12462132	placenta_normal

FIGURE 1B

```

1 MDFLLALVLV SSLYLQAAAE FDGRWPRQIV SSIGLCRYGG RIDCCWGWAR
51 QSWGQCQPFY VLRQRIARIR CQLKAVCQPR CKHGECIGPN KCKCHPGYAG
101 KTCNQDEHIP APLDQGSSEQP LFQPLDHQAT SLPSRDLNEC GLKPRPCKHR
151 CMNTYGSYKC YCLNGYMLMP DGSCSSALTC SMANCQYGCD VVKGQIRCQC
201 PSPGLQLAPD GRTCVDVDEC ATGRASCPRF RQCVNTFGSY ICKCHKGFDL
251 MYIGGKYQCH DIDECSLGQY QCSSFARCYN VRGSYKCKCK EGYQGDGLTC
301 VYIPKVMIEP SGPIHVPKGN GTILKGDTGN NNWIPDVGST WWPPKTPYIP
351 PIITNRPTSK PTTRPTPKPT PIPTPPPPPP LPTELRTPLP PTTPERPTTG
401 LTTIAPAAST PPGGITVDNR VQTDPOKPRG DVFIPRQPSN DLFEIFEIER
451 GVSADDEAKD DPGVLVHSCN FDHGLCGWIR EKDNDLHWEP IRDPAGGQYL
501 TVSAAKAPGG KAARLVLP LG RLMHSGDLCL SFRHKVTGLH SGTLQVFVRK
551 HGAHGAALWG RNGGHGWRQT QITLRGADIK SVVFKGEKRR GHTGEIGLDD
601 VSLKKGHCSE ER (SEQ ID NO:2)

```

FEATURES:

Functional domains and key regions:

Prosite results:

PDOC00001 PS00001 ASN_GLYCOSYLATION

N-glycosylation site

320-323 NGTI

PDOC00005 PS00005 PKC_PHOSPHO_SITE

Protein kinase C phosphorylation site

Number of matches: 10

1	157-159	SYK
2	222-224	TGR
3	284-286	SYK
4	354-356	TNR
5	358-360	TSK
6	362-364	TTR
7	366-368	TPK
8	531-533	SFR
9	573-575	TLR
10	602-604	SLK

PDOC00006 PS00006 CK2_PHOSPHO_SITE

Casein kinase II phosphorylation site

Number of matches: 3

1	213-216	TCVD
2	392-395	TTPE
3	453-456	SADD

PDOC00007 PS00007 TYR_PHOSPHO_SITE

Tyrosine kinase phosphorylation site

246-252 KGFDLMY

PDOC00008 PS00008 MYRISTYL

N-myristoylation site

Number of matches: 10

1	156-161	GSYKCY
2	172-177	GSCSSA
3	194-199	GQIRCQ
4	238-243	GSYICK
5	268-273	GQYQCS
6	283-288	GSYKCK
7	497-502	GQYLTV
8	509-514	GGKAAR
9	538-543	GLHSGT
10	552-557	GAHGAA

FIGURE 2A

PDOC00010 PS00010 ASX_HYDROXYL
 Aspartic acid and asparagine hydroxylation site
 Number of matches: 3

1	151-162	CMNTYGSYKCYC
2	233-244	CVNTFGSYICKC
3	278-289	CYNVRGSYKCKC

PDOC00016 PS00016 RGD
 Cell attachment sequence

429-431	RGD
---------	-----

PDOC00021 PS00022 EGF_1
 EGF-like domain signature 1

92-103	CKCHPGYAGKTC
--------	--------------

PDOC00021 PS01186 EGF_2
 EGF-like domain signature 2
 Number of matches: 3

1	92-103	CKCHPGYAGKTC
2	160-174	CYCLNGYMLMPDGSC
3	287-300	CKCKEGYQGDGLTC

PDOC00913 PS01187 EGF_CA
 Calcium-binding EGF-like domain signature
 Number of matches: 3

1	136-160	DLNECGLKPRPCKHRCMNTYGSYKC
2	216-242	DVDECATGRASCPFRQCVCNTFGSYIC
3	261-287	DIDECSLGQYQCSSFARCYNVRGSYKC

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	3	23	0.942	Putative
2	163	183	0.765	Putative
3	397	417	1.209	Certain

SignalP results:

< Is the sequence a signal peptide?

#	Measure	Position	Value	Cutoff	Conclusion
	max. C	20	0.657	0.37	YES
	max. Y	20	0.713	0.34	YES
	max. S	6	0.960	0.88	YES
	mean S	1-19	0.901	0.48	YES

Most likely cleavage site between pos. 19 and 20: AAA-EF

FIGURE 2B

BLAST Alignment to Top Hit:

>CRA|50000098943299 /altid=gi|15795193 /def=ref|NP_277060.1|
nephronectin [Mus musculus] /org=Mus musculus
/taxon=10090 /div=ROD /dataset=nraa /length=609
Length = 609

Score = 1190 bits (3045), Expect = 0.0

Identities = 536/610 (87%), Positives = 569/610 (92%), Gaps = 3/610 (0%)

Query: 1 MDFLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCGWARQSWGQCQPFY 60
M LLA VL SLYLQ AA+FDGRWPRQIVSSIGLCRYGGRIDCCGWARQSWGQCQPFY
Sbjct: 1 MAVLLAAVLASSLYLQVAADFGRWPRQIVSSIGLCRYGGRIDCCGWARQSWGQCQPFY 60

Query: 61 VLRQRIARIRCQLKAVCQPCKKHGECIGPNKCKCHPGYAGKTCNQDEHI-PAPLDQGSEQ 119
VLRQR+ARIRCQLKAVCQP+CKKHGEC+GPNKCKCHPG+AGKTCNQDE P PLDQGSEQ
Sbjct: 61 VLRQLARIRCQLKAVCQPQCKKHGECVGPKNKCKCHPGFAGKTCNQDESFHPTPLDQGSEQ 120

Query: 120 PLFQPLDHQATSLPSRDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALT 179
PLFQP DHQAT++PSRDLNECGLKPRPCKHRCMNT+GSYKCYCLNGYML+PDGSCSSAL+
Sbjct: 121 PLFQPPDHQATNVPSRDLNECGLKPRPCKHRCMNTFGSYKCYCLNGYMLLPDGSCSSALS 180

Query: 180 CSMANCQYGCDDVVKQIRQCQPSPLQLAPDGRTCVDVDECATGRASCPFRQCVNTFGS 239
CSMANCQYGCDDVVKGQ+RCQCPSPGLQLAPDGRTCVD+DECATGR SCPRFRQCVNTFGS
Sbjct: 181 CSMANCQYGCDDVVKGVRCQCPSPGLQLAPDGRTCVDIDEATGRVSCPRFRQCVNTFGS 240

Query: 240 YICKCHKGFGLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKKEGYQGDLT 299
YICKCH GFDLMYIGGKYQCHDIDECSLGQ+QCSS+ARCYN+ GSYKC+C++GY+GDGL
Sbjct: 241 YICKCHTGFGLMYIGGKYQCHDIDECSLGQHQCSSYARCYNIHGSYKQCRDGYEGDGLN 300

Query: 300 CVYIPKVMIEPSGPIHVPKNGTILKGDITGNNNWIPDVGSTWWPPKTPYIPPIITNRPTS 359
CVYIPKVMIEPSGPIH+P+ NGTI KGD G+ N IPD GST WP KTPYIPP+ITNRPTS
Sbjct: 301 CVYIPKVMIEPSGPIHMPERNGTISKDGGHANRIPDAGSTRWPLKTPYIPPVITNRPTS 360

Query: 360 KPTRRPTPKPTPIPTPPPPPLPTELRT-PLPPTTPERPTTGLTTIAPAASTPPGGITVD 418
KPTRRPTP PTP TTPPPPPPLPTE RT PLPPT PERP+T TTIAPA ST ITVD
Sbjct: 361 KPTRRPTPNPTPQPTPPPPPLPTEPRTTLPPT-PERPSTRPTTIAPATSTTRVITVD 419

Query: 419 NRVTQDPQKPRGDVFIPRQPSNDLFEIFEIERGVSADEAKDDPGVLVHSCNFDHGLCGW 478
NR+QTDPQKPRGDVFIPRQP+NDLFEIFEIERGVSADE KDDPG+L+HSCNFDHGLCGW
Sbjct: 420 NRIQTDPQKPRGDVFIPRQPTNDLFEIFEIERGVSADEEVKDDPGILHSCNFDHGLCGW 479

Query: 479 IREKNDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPGLRLMHSGDLCLSFRHKVTG 538
IREKD+DLHWE RDPAGGQYLTVSAAKAPGGKAARLVL LG LMHSGDLCLSFRHKVTG
Sbjct: 480 IREKSDLHWETARDPAGGQYLTVSAAKAPGGKAARLVRLGLHMHSGDLCLSFRHKVTG 539

Query: 539 LHSCTLQVFVRKHGAHGAALWGRNGGHWGRQTQITLRGADIKSVVFKGEKRRGHTGEIGL 598
LHSCTLQVFVRKHG HGAALWGRNGGHWGRQTQITLRGAD+KSV+FKGEKRRGHTGEIGL
Sbjct: 540 LHSCTLQVFVRKHGTHGAALWGRNGGHWGRQTQITLRGADVKSIVFKGEKRRGHTGEIGL 599

Query: 599 DDVSLKKGHC 608
DDVSLK+G C
Sbjct: 600 DDVSLKRGRC 609 (SEQ ID NO:4)

FIGURE 2C

```

>CRA|46000i03872918 /altid=gi|15128103 /def=gb|AAK84391.1|AF397007_1
  (AF397007) nephronectin [Mus musculus] /org=Mus musculus
  /taxon=10090 /div=ROD /dataset=nraa /length=592
  Length = 592

Score = 1147 bits (2934), Expect = 0.0
Identities = 520/610 (85%), Positives = 552/610 (90%), Gaps = 20/610 (3%)

Query: 1  MDFLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY 60
      M  LLA VL SSLYLQ AA+FDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPF
Sbjct: 1  MAVLLAAVLASSLYLQVAADFGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPF-- 58

Query: 61  VLRQRIARIRCQLKAVCQPRCKHGEICIGPNKCKCHPGYAGKTCNQDEHI-PAPLDQGSEQ 119
      VCQP+CKHGEIC+GPNKCKCHPG+AGKTCNQDE  P PLDQGSEQ
Sbjct: 59  -----VCQPQCKHGEICVGNKCKCHPGFAGKTCNQDESFHPTPLDQGSEQ 103

Query: 120  PLFQPLDHQATSLPSRDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALT 179
      PLFQP DHQAT++PSRDLNECGLKPRPCKHRCMNT+GSYKCYCLNGYML+PDGSCSSAL+
Sbjct: 104  PLFQPPDHQATNVPSRDLNECGLKPRPCKHRCMNTFGSYKCYCLNGYMLLPDGSCSSALS 163

Query: 180  CSMANCQYGCDDVVKQIRQCQPSGLQLAPDGRTCDVDDECATGRASCPFRQCVNTFGS 239
      CSMANCQYGCDDVVKQ+RCQCQPSGLQLAPDGRTCDV+DECATGR SCPRFRQCVNTFGS
Sbjct: 164  CSMANCQYGCDDVVKQVRCQCPSPGLQLAPDGRTCDVIDECATGRVSCPRFRQCVNTFGS 223

Query: 240  YICKCHKGFIDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKKEGYQGDLT 299
      YICKCH GFDLMYIGGKYQCHDIDECSLGQ+QCSS+ARCYN+ GSYKC+C++GY+GDGL
Sbjct: 224  YICKCHTGFDLMYIGGKYQCHDIDECSLGQHQCSSYARCYNIHGSYKQCRDGYEGDGLN 283

Query: 300  CVYIPKVMIEPSGPIHVPKNGTILKGDGTGNNNWIPDVGSTWWPPKTPYIPPIITNRPTS 359
      CVYIPKVMIEPSGPIH+P+ NGTI KGD G+ N IPD GST WP KTPYIPP+ITNRPTS
Sbjct: 284  CVYIPKVMIEPSGPIHMPERNGTISKDGGHANRIPDAGSTRWPLKTPYIPPVITNRPTS 343

Query: 360  KPTTRPTPKPTPIPTPPPPPLPTELRT-PLPPTTPERPTTGLTTIAPAASPPGGITVD 418
      KPTTRPT PTP PTPPPPPPLPTE RT PLPPT PERP+T TTIAPA ST ITVD
Sbjct: 344  KPTTRPTPNPTPQPTPPPPPLPTEPRTTLPPT-PERPSTRPTTIAPATSTTTRVITVD 402

Query: 419  NRVTQDPQKPRGDVFIPRQPSNDLFEIFEIERGVSADDEAKDDPGVLVHSCNFDHGLCGW 478
      NR+QTDPQKPRGDVFIPRQP+NDLFEIFEIERGVSAD+E KDDPG+L+HSCNFDHGLCGW
Sbjct: 403  NRIQTDPQKPRGDVFIPRQPTNDLFEIFEIERGVSADDEEVKDDPGILHSCNFDHGLCGW 462

Query: 479  IREKNDNLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPLGRMLHSGDLCLSFHRHKVTG 538
      IREKD+DLHWE RDPAGGQYLTVSAAKAPGGKAARLVL LG LMHSGDLCLSFHRHKVTG
Sbjct: 463  IREKSDNLHWEIARDPAGGQYLTVSAAKAPGGKAARLVLRLGHLHSGDLCLSFHRHKVTG 522

Query: 539  LHSCTLQVFVRKHGAHGAALWGRNGGHWGRTQITLRGADIKSVVFKGEKRRGHTGEIGL 598
      LHSCTLQVFVRKHG HGAALWGRNGGHWGRTQITLRGAD+KSV+FKGEKRRGHTGEIGL
Sbjct: 523  LHSCTLQVFVRKHGTHGAALWGRNGGHWGRTQITLRGADVKSIVFKGEKRRGHTGEIGL 582

Query: 599  DDVSLKKGHC 608
      DDVSLK+G C
Sbjct: 583  DDVSLKRGRC 592 (SEQ ID NO:5)

```

FIGURE 2D

```

>CRA|84000042916333 /altid=gi|15430248 /def=gb|AAK96011.1|
      (AY035899) nephronectin long isoform [Mus musculus]
      /org=Mus musculus /taxon=10090 /div=ROD /dataset=nraa
      /length=578
      Length = 578

      Score = 1128 bits (2885), Expect = 0.0
      Identities = 512/609 (84%), Positives = 543/609 (89%), Gaps = 32/609 (5%)

Query: 1  MDLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY 60
      M LLA VL SLYLQ AA+FDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY
Sbjct: 1  MAVLLAAVLASSLYLQVAADFGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY 60

Query: 61  VLRQRIARIRCQLKAVCQPCKKHGECIGPNKCKCHPGYAGKTCNQDEHIPAPLDQGSEQP 120
      VLRQR+ARIRCQLKAVCQP+CKKHGEC+GPNKCKCHPG+AGKTCNQ
Sbjct: 61  VLRQLARIRCQLKAVCQPQCKKHGECVGPKNKCKCHPGFAGKTCNQ----- 105

Query: 121  LFQPLDHQATSLPSRDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGCSSALTC 180
      DLNECGLKPRPCKHRCMNT+GSYKCYCLNGYML+PDGSCSSAL+C
Sbjct: 106 -----DLNECGLKPRPCKHRCMNTFGSYKCYCLNGYMLLPDGCSSALSC 150

Query: 181  SMANCQYGC DVVKGQIRCQCPSPGLQLAPDGRTCDVDDECATGRASCPRFRQCVNTFGSY 240
      SMANCQYGC DVVKGQ+RCQCPSPGLQLAPDGRTCDV+DECATGR SCPRFRQCVNTFGSY
Sbjct: 151  SMANCQYGC DVVKGQVRCQCPSPGLQLAPDGRTCDVIDECATGRVSCPRFRQCVNTFGSY 210

Query: 241  ICKCHKGFDMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRSYKCKCKEGYQGDGLTC 300
      ICKCH GFDLMYIGGKYQCHDIDECSLGQ+QCSS+ARCYN+ GSYKC+C++GY+GDGL C
Sbjct: 211  ICKCHTGFDLMYIGGKYQCHDIDECSLGQHQCSSYARCYNIHGSYKQCQRDGYEGDGLNC 270

Query: 301  VYIPKVMIEPSGPIHVPKNGTILKGDGTGNNNWIPDVGSTWWPPKTPYIPIITNRPTSK 360
      VYIPKVMIEPSGPIH+P+ NGTI KGD G+ N IPD GST WP KTPYIPP+ITNRPTSK
Sbjct: 271  VYIPKVMIEPSGPIHMPERNGTISKGDGGHANRIPDAGSTRWPLKTPYIPPVITNRPTSK 330

Query: 361  PTTRPTPKPTPIPTPPPPPLPTELRL-TPLPPTPERPTTGLTTIAPAASTPPGGITVDN 419
      PTTRPTP PTP PTPPPPPLPTE R TPLPP TPERP+T TTIAPA ST ITVDN
Sbjct: 331  PTTRPTPNPTPQPTPPPPPLPTEPRTTPLPP-TPERPSTRPTTIAPATSTTTRVITVDN 389

Query: 420  RVQTDQPQKPRGDVFIPIRQPSNDLFEIFEIERGVSADDEAKDDPGVLVHSCNFDHGLCGWI 479
      R+QTDQPQKPRGDVFIPIRQP+NDLFEIFEIERGVSAD+E KDDPG+L+HSCNFDHGLCGWI
Sbjct: 390  RIQTDQPQKPRGDVFIPIRQPTNDLFEIFEIERGVSADDEVKDDPGILIHSCNFDHGLCGWI 449

Query: 480  REKDNDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPLGRMLHSGDLCLSFHVKVTGL 539
      REKD+DLHWE RDPAGGQYLTVSAAKAPGGKAARLVL LG LMHSGDLCLSFHVKVTGL
Sbjct: 450  REKSDLHWETARDPAGGQYLTVSAAKAPGGKAARLVLRLGHLHSGDLCLSFHVKVTGL 509

Query: 540  HSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIKSVVFKGEKRRGHTGEIGLD 599
      HSGTLQVFVRKHG HGAALWGRNGGHGWRQTQITLRGAD+KSV+FKGEKRRGHTGEIGLD
Sbjct: 510  HSGTLQVFVRKHGTHGAALWGRNGGHGWRQTQITLRGADVKSIVFKGEKRRGHTGEIGLD 569

Query: 600  DVSLKKGHC 608
      DVSLK+G C
Sbjct: 570  DVSLKRGRC 578 (SEQ ID NO:6)

```

FIGURE 2E

Hmmer search results (Pfam):

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
CE00409	E00409 fibulin	108.3	3.8e-30	3
PF00629	MAM domain.	88.2	1.7e-22	1
PF00008	EGF-like domain	68.0	1.7e-17	5
CE00283	E00283 selectin	13.0	0.014	2
PF01278	Omptin family	1.9	6.6	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00008	1/5	77	103 ..	1	45 []	17.3	0.0036
CE00283	1/2	74	105 ..	119	153 ..	10.7	0.059
PF00008	2/5	140	172 ..	1	42 [.	19.4	0.00091
PF00008	3/5	180	214 ..	1	45 []	19.0	0.0012
CE00409	1/3	136	227 ..	415	514 ..	69.4	4.5e-19
CE00283	2/2	226	248 ..	124	146 ..	8.7	0.2
PF00008	4/5	220	249 ..	1	41 [.	10.8	0.25
CE00409	2/3	232	273 ..	429	473 ..	35.2	2.3e-09
CE00409	3/3	277	298 ..	344	365 ..	16.3	0.00053
PF00008	5/5	265	300 ..	1	45 []	31.7	2.9e-07
PF01278	1/1	477	493 ..	241	257 ..	1.9	6.6
PF00629	1/1	469	610 ..	1	170 []	88.2	1.7e-22

FIGURE 2F

1 GAGAAAATTG AGATTACTAC CTGCAAGGTG TCATTACCTG GTAAGAAGCC
51 TATCAAAAGT TTGTCCTCCT GAAAAAGTAG TTATTGCTAA AAGCTAGCTG
101 TTTTGATCTC ATTCTTGCTC ATTTGTTTTT AAGACTGAGA TAATGAAATG
151 TCACTCCCAT GGCAACTCTG CCTCTTTTTC GGAATGATCA TTGGTGGTCA
201 TAGTTGCAGC ATAATAACCA GTTAGACCTT GGAAATCCTT TAGATTCTCC
251 TTATTCCATG ATTTAACAAA GACTGATATA ATTAGCTACA TTTTACTGAA
301 GGGAGAAGCT AAAGTTCACA GGCAGAATTC AATTTAATCC AATCCATCTG
351 TTTNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNCT ATTGGTTAGA
451 CTAATGTCTC ACGGGACCCA GAGTGGTAGG GAGGCAATGA CAACACAATA
501 CATTAGAGC TACCTTAGAG CATGCTAAGT GTAATAGAAG TATATAAAAC
551 ACTAGTCTCA GCCAACCAAA AAGTCAGAAA AGGCTTCCTG GGGGACTGTC
601 GGCTAAATTT AAACCTCAAAG GGTAAGGGGA GATTGTCCAA ATGAAGAAGG
651 AAGAGGAATG AAGAGGAGGC TAAATTTAAA TGAAGAAGGA GTGTTTTCTA
701 GGCCAAAGCA AATATATGGA AAACAAAAGA AAGACGCATG TAGATGGGGC
751 ATTACACTTT TCCCTCCAGT TATTTATCCT GTTTTCATCC ACCACTCTTC
801 GTCTTTTCTC TAGATCTCCA GTTTTTTAGC CGTATATTAC CCCCTTTCTC
851 TCTAATCATC CATGTCACAC AGTGAGGTTT ATTTGTAAAT CTAACCCAGG
901 CACCAACTAA CCAACCAACA AACAAAAACA GCTAAGGAGT AGCCACTGGA
951 ACCTGGAACC ATGCCTACAC TTACAAAAAT TTGATTTTCT GCAGAAAAAA
1001 TATTTTGACT CCTACATTTT TGGACTTCAT TAGAAGGACC TGAAATGGAT
1051 GACACCAAGC TGTTTGCTTA AAATAATGTC CCAAGCCTGA ATTGGCATGG
1101 ATCTTTCTTG AGATTAAAAAT AGAACTTGT TTTGCTAACT GAAAACAACCT
1151 TAGAAATCAA AGAGCCATTT AAGTTGAAAC CATTATTTT CCTTTCCTTG
1201 AAGAAAATTC CTGTTTTTAC ACACACTGAA TGATCAGGAT AGTGAATCAC
1251 CCTACCACAG AACTTTCAT TAAAAATTTG AAGTTGTAGA AACCTCAAAA
1301 AGAAAATGAA GATGGGGGGA AAACGTTTGT AATGTAGCAA ATGANNNNNN
1351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNAGAT ATCACTGCAC
1901 AAGGCTGTCA TTCGAGGTAA GGTGGGAAG AGGGGTGACT TGATTGGGCT
1951 CATGTACATC TTAGGATAAT CATGTTTTCA TTTGAACACA GTCATTTAGT
2001 ACAATACTTT TTA CTCTGCA ACAACTGAAC AATGATTACT TAAAATGGTT
2051 TCTAAGTGAT CCCCTCAAGA TCAGGCACCA CAACACTACT AGGCAGTAGC
2101 AGCCCTTCCT CTCTGGCAGG CACTCATCCC TAGAGGGGCA GTCCTGTCTC
2151 AATGTGCTCG GGAGTTGAGC TGATAGAGAC AGAAACAGTC TGCTCATTTT
2201 TTCTTCAAAA CCGGCTGGA GAGCCAAATA TTTCTTTTGG TTAGTTACAA
2251 AATAAAAAAT ATGCCTTTGC CAACTCTGCG GGAAAAGCAA ACTCCCAGAG
2301 TTTGCCTTCG CTAATTTGTC CAAGTCGTGC TGTGGTTTTA GTAAATGCAG
2351 AACTGCTCA GCTCCAGCCC CATAAACCTC TCTGCTCTAG GGCTTCTCCG
2401 CCTCCGGTG AAAGCTACTG CTCGCCCTG CAGTCACCAC CTGTTCCGGC
2451 GGAACCTGCG GAGCGTGCAC CTACGCCTCG GGCTCCTTTC CTCCTCCACT
2501 CCCCTTTCTC GCTGGGCACC CTGCTTTCCC TCTCCCAGAG AGGGTTTGCA
2551 ACTTTTCTCC CAGGCTGGGG CTCGCCCTGC TTGGCTAACC CCAAGAGCC
2601 ACTGCCGTPC CGCAGCGCCC CTGCCCCGA GTTGCTGCC CCGCTGGGCC
2651 CCCGGGAGGA GCGGAGCGCG CTCACCTTC GCCCGGGGCT GGGAGGGCGG
2701 CGAGTCGGGC GCACGCGCAC CCCCTGCCCG CCCCTGGCGC CCTCCCCGC
2751 GGGCGGTGCA GCTACCCCTG CAGCGCTCC CTAAGCTAGA AGGGAGCGGG
2801 AGGGGGCTCC GGGCGCGCG CAGCAGACCT GCTCCGGCG CGCGCCTCGC
2851 CGCTGTCTC CGGAGCGGC AGCAGTAGCC CGGGCGGCG GGGCTGGGGG
2901 TTCCTCGAGA CTCTCAGAGG GCGCCTCCC ATCGGCGCCC ACCACCCCAA
2951 CCTGTTCTC GCGCGCCACT GCGCTGCGCC CCAGGACCCG CTGCCCAACA

FIGURE 3A

3001 TGGATTTTCT CCTGGCGCTG GTGCTGGTAT CCTCGCTCTA CCTGCAGGCG
3051 GCCGCCGAGT TCGACGGGAG GTGAGCTGGG CCCCAGGGCG CCCTCTCCTC
3101 CTTCCCGCGC TAATTTACACA CTCACCTGTCT TGGGTCACTT TTCCCCCGCG
3151 GGTTCGTGG TCAGAGAGGC GTCTCCTCCA TCCAGAAGTT GGGCCACCGC
3201 ACAGCGTGGC GCGAGGAGAG CGGTCCAGCG GCTCCGAGTG CCGCCCCGAG
3251 GCGGAGAGGG CGCGCCCTTG CGAGTCTGGG ACCCCATCCG CGGCCCCCG
3301 AGGGCGACTC GCCCCGCTC GGGAAATTAGG ACTGAGGGAG AGGAGCCGCT
3351 GGAGCCTGGG ATCTCGGCTC TGAGGGCGCG GTTTAGCCAC CTACGCCGAG
3401 GTGACGCGCG AAACATCCCT TACCCGGGAA ACTCCCGCGC CTGAAC TAGA
3451 CGGCTCTTCA CTGGGGAAGC TTCCAGGCCC CCGGGGGGAG GCCCGGGCTC
3501 TGCTCAGGGC TCTCGGGGCC GCTCACACAG AGAGTGGGTG CGAGTCAGCG
3551 ACTGGGCTAC TGTGGGCTC TCCATTTGGT TTTCTTGAGG
3601 GAAGGAGACT CAAAATGAGG ACCGAGGGT GGGCGCTCCG TGAATGTGAG
3651 CATGAGTGTG TGGATGTGTG TGTGAGAGCG CACACACTGC GCCGCTCCTC
3701 AGACTCGGGC GAGCCTGACG GCGGCGTGCT GTGACAGGTT CCAACAACCT
3751 CGGGCCGCGT CTCCGCTGTC ACTCAGCCGG TCCTCCCGCG CTCCGGGGCC
3801 GCTCCGGTGT GTGAGAGACA CTGGGTCTGT CGGGAGGGTG TGCTCGGTCC
3851 CCCTCACCTC TTGCAATTA CAGACTAGGC TCGTCCCGG TGCAGATGGC
3901 TGCTGCGAAA AGAGGTTTA CTCTGGCGCA CACCGTCGCC CGGTGCGCTT
3951 CCTTCAGCGA CCTCTGCCCC CACCCCTCCG TGTAGGTCCC TGCCTAGTCC
4001 TAAAGAAAGA TGCCGCACCT GTTTTACCTT TAATCTTGGA ACGAATCAAT
4051 TCCACAATTG ATTGCTTTC TTGTACCCGG AGGCGAACGG AGGGGAACCA
4101 GTGGAGCGCC AAGAAAAAGC ACAGCTGCTT GTTCACTACC TGGTCGGAGG
4151 AAATAGAAAG TTCGGGGTTA TTTTGTACTT TGGGCCTGGG GGTAGAAAGG
4201 CAGGTAAAAG AAAAGGGGAA TTGAAAAAGA TAAGGAGACT TTAAGAAAAC
4251 GTGATAGCAG CCAGAGTGTA GACCTTTTTA TTTTATTTT AAAAGCAATT
4301 CTGTGCTCAC ATTTGGGTAT GTTATGCATT TCTTGACAC ATATTTGCAA
4351 CAAGAAACCC ATCACAAGAT ATGCATATGA TGTGAATGCA TATAGCTTTT
4401 TGTAACTTTA AAAAGATGTC TAAAACAGCA AATTAAAATG TTGATACCCA
4451 GAGGAAAAAA GTCATTTAGG TGTGACTCTA AACAAGGAAA CAATTTAGCA
4501 AATAATGTGT CAACGTGTA TCAGGTGGAT AAATCTGTAT CCTGAAATTA
4551 CTTTCTTTAG GCATTATTTA CATTAGAGAG GAAAATACTA TGAATTGTTG
4601 ATCTAAGCAC GTTTCAAACC ACAAGGATAC TTGAGATATC AGCTACACTT
4651 TAACTCTCC TTGGTTTTAT TTAGACTTTT TTCTAGTACT TTTTGTTTT
4701 TTTTGTCTCT GCCACAGTTA TGTCTCAAAA AGAGCTCTGT TACACTTGAT
4751 TTTTGAGAAA CCTTCTGCT CCTTCCCCCA TCCACCTTTT ATATGAGGCA
4801 GAAATTTTTT CTGCTTTGAC ATGTCTTACT TAATACTTTT CAGTTTATGG
4851 TGAATCAGTC AACCTGGCT TTCACCCAG TAACAGGTGT GGGTCTTTGA
4901 GGAATCTGTT TTATTTCAAG CTTCAGAAAC TATCCTGTGA GTGGCAGCCT
4951 TGAGAATTGT TGGTGTAGC TGGTGCTATT TCTACTTAAG AAAGCATCTT
5001 TCCCTCCCGT TTTTCTCCT ACCCTGTGT GTTGGGGGAT GGGGTATTTA
5051 ACTGTCTGAA ATTTGACATT AGATCATAGC AGAACTATTT CTGTGAAGGT
5101 GTTTTAACCT TAAATTTCTG CAGGCATAAA AGAGTTTGTA GAAAATTGTT
5151 GGGTGTGGTT GACATTTTTT CTGAAGTAAT AATACAGAGG AAATTACCTT
5201 TCTTCTCAA AGTGTCTTA TTTATTTTGA ATCTTTTTT TTGGTAGGTG
5251 CCCCAGGCAA ATAGTGCAT CGATTGGCCT ATGTCGTTAT GGTGGGAGGA
5301 TTGACTGCTG CTGGGGCTGG GCTCGCCAGT CTTGGGGACA GTGTCAGCGT
5351 GAGTATCAAG CCTGGGGACT TCAGTTCCCT GGGAGGTGTG GCTTTCCACC
5401 TTGTTTATGG CTTACCCCCA CATATCAGAG GGTTCATTAC TGAGCAAGGC
5451 TTGGCCTTGC AGGTCGACT TGGGGATTTT CAGGTACAGT CCAGACTCCT
5501 TATTCTGCTT CTTTTAGCT TTAGCCACCT GTATTACGGC CCAGCTTTGT
5551 CATTACAGAG GAGCCACCTT AAATGTTTCT TTAGCATTTT TTCCCCCAA
5601 ATCTGCATCT CCTTCTGTT TTCTTACCAG GTTTATTNNN NNNNNNNNN
5651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5701 NNNNNNNNNN NNNNNNNNNN NNNNNNNCTGA CTCGCTTGCT CCAGGCTGGG
5751 AGATGGGGAT AAGATAATGA GTCTAGACCG TATTTACACT GGGGATAAGA
5801 TAATGAGTCT AGACCGTATT TACACAAAAT AAACCTCTGAA CCTTAGAGTA
5851 CTTATATAAT CCTCACATA GCTGTTGTGT AAAGTGGCCC ATTTGTAAT
5901 CCTTTTTTGG ACTGGGGTTG TATTGACAGT GTTTTAAATG GACCACCCAG
5951 TATAATGAAA CAAAGCCACT GAAACAACAT TATTTGTAAG TTCTATAAGT

FIGURE 3B

6001 AACCAACCTC ATGTAATAAG TACCAAATCA GAAAGCTGAT TTTCACCTTT
6051 TCTGTGACCA CAAATGGATA TTTTAATGGT AAAATTTAGA GCTACATCAA
6101 AAGAGTTTGA GAAATTTAGG AAATTTGTCC CCAGCATTTT ATCGCGATCT
6151 TAAAATTGTA TCTCACTCCT ACTGCAAAAA AATAGTCTTC AAATGATCAA
6201 GTACCTTTCC AGAGCACCCCT TAGAGTGCTT GGGGCGGGGG TAGGGGGCAC
6251 TTCTGGTAAG ATGATGGGAA CTAAGTTGGG TTCTACATTG GGATATATAT
6301 TTTATTGCTA ATGAGGAGGA GGCTTAGAGG AAGAGAGAAG GGCAGTTACG
6351 AAGGCTAGAG CTGGCAATGG AGAAGCCTGC CTTAGAGATG GGTGCTAGT
6401 GTGAGGAGTC AGGCAAATTT AAGTTCAGGA AAGTTAGGAG TTCCCTCTGC
6451 TATTTTAATT TTTGAGGATG CTTGCAATGT CTTCTTAAT TTTGTGAAAG
6501 AGGGACAGTG ACAGTCACAG ATTGACTCTA ATTGCACATA AAGACCACAA
6551 TCTCTGGTTG GGAATAGAAA GGTAAAGGAA ATGAATGTTT GCCTACCTGG
6601 TATGGAATTT GAGAACCAAC AGATTCTAAT AACCAAAATG TGAAGAAAGG
6651 ACCCTTCTGT TGGCCCAACA CACCTACACA TAACCCTCCT GAGTGAAAAA
6701 TGAGTAGTTC TATACCTGCA GTCTCCAGCT GTGCAAATAC TTCTGATACT
6751 ACAGAAGACT AAATTCACC AGGCACCATT CTCTTTTTTT GATCATCTTC
6801 CCTTAAACA ATATTGAATA GACTAACCAG TGAGTGACATA GCAGCTTTGC
6851 CTACTTCTTT TGTTTACTGG AAAGTGGAGT TACCACCATC TCCCTTTAAC
6901 AGAATGTAAT TGACCCCGG TGCAAAGGT CTAGCCAAGC AGCTCGCACA
6951 TAGCAGTTAG GCAAACATTC GTTTGCTTCA CCTTTTATGT TTATACTTCC
7001 AAAATCAAAG TAGTTCCAAG GTTTTCTATG TTACAGTGAA ATAAAATCCC
7051 TTTCAATTAA AAAGGCACAA ATGGTTCTTT ACTATTAACA TTGAAGTTGG
7101 TGAGGTTGTA AAAATTAGCT CAAAGGTGAA TGTTCCTTC TGTGTATTTT
7151 ATTTTTCAG CATCTTAGAT GGCGCAAATG TCTCTGTCA GTTCAGAGTT
7201 CTGCCTGTCT TTGTTTGTAT ATAAGCAGGT AGAGGAATGT GGGGCTGAGA
7251 AGTAAGCTTG AAGGGGCAGA ACAAACCAA AAGAGGCTGA TCAGATTGAA
7301 TGAAATATCT CTGAAAACCTC TTGATTATTT TAAAGAAAGT CTTTATGAAA
7351 TTAAGGTTT TGTCCTACG TTTGTTCAAG AAAATGCCTT GCTATTGTAT
7401 AACCAATCAA TCTAATATGA TTCCTTATAA AGATTCCAAA GAACCTCTAG
7451 CGAATTTAAT GTGAGAAATG TTTTTCCTT TCGACCTTT AGATAATCAT
7501 GTAGTTCTTT CCCATAAGGA AGGGCTATTC TCCCTTCCTC ATCAGAGGTG
7551 CTCTGGTTCT CTTCTTTCT GAATGATTCA CTTTGAATT TTCCTTCAA
7601 ACAGCATAGC AAAACAAAAA GAACTATTC CCATTACTGC ATAGATCTTC
7651 CCAACTTATC CCATAGAAGG TGTGTTTGTA GGTAGGAAA GGTGGTGGTA
7701 CCCTCATAT ATTATTAAAC AGACCTTATG CCTTAGTGCA GTGACCTTT
7751 GGTGAGTGTT TATTCGGTTT GGTGAGCAT TTGTGAGATG TTTGACTTAA
7801 TATTCATGTG AGTCAAATGT ACGTATCTGT ACAAACATG CTGCCCTTCA
7851 TTTTCACTTG CTATCTTCT ATCCATGGTC TTGTTGGAG AAACCGACTA
7901 ATGTTGCAGG ATGCTAAAGC TGGTAGACCT CTCCTTCTGG CTCACTATGT
7951 CTAAGCAGAG CCAGATATAG CTGGGAAACT TTATATCCTT CCCTCTAGGA
8001 CTCAGGAAGG AAGGATCAGA GATGCTACTC AAATGGGCAT AGAACCTGTC
8051 CTGCTGCTTC CTGCCTGTAC CACCCAACAT TCTATCCAA CATTCTGCC
8101 TGCTAGGGAA AGCATGACCC GTTCCAGAAA TAGGCAGGTT GTGTTTTCAT
8151 AGTTCTTTTG TAGGTCATAC CCTTGGTTGN NNNNNNNNNN NNNNNNNNNN
8201 ATTCTTTGTT TTGTGTTCTT GACTCTTAAA CAGATCTCTA GCATATTGAA
8251 AATTCAACAT TTGATTTTCT AACTGTCATG GGCTTTACTT TTATTGACTA
8301 TTGATGTGGC TGTTTATTGT CAGGTGAAAA TTTTTTAATA TGTTACACACA
8351 CTGATGATTG CATATTTGCA GCACACAGCA TCTTAAACCA CTCAGAGGTT
8401 TGTCACAAAA TGTGTGTTTC TTGCTCTGTA ATTTTTTTGT CATTTTGATG
8451 GCATATTTTA ATTATGCTTT TATTTCTCTC CTTCTACCAA GTGGAGACCT
8501 CTGTAACAAA GATTTTTCAG GGTGTCATAT TTCATCTTTA AAAGCTGTGT
8551 AACTGTGGA AACAAATGTT AACCAAAAGC ACTAGATAAT TTAGGATAAC
8601 TGAGTTCAAG TTCTGACTTT TGTACCTGAA TAGGTGTCTA ACATTTCTCA
8651 GCTACAGTTT TCTTATTACA CAAGCATATT TCCAAGGTAT GTTGAAGCTC
8701 TAAACTGCAA CTGAAAACCTT CCTTAATGTA GAAAACTAT ATAGGATCTA
8751 AATATTGTAT TTTTGTATGG TTGTCCTTCT GTTTACTCGG AGATTTGACT
8801 GTATATGTGG CTTATGACAA TAGCATTTT GTTAAAAGCA TTTTATAGAA
8851 GTGTGAAGAA AAACATAAAA TACAACCAGT TCCAAGGTTT AACAAAACCTA
8901 TTCCGTTTCT GAGTTCCTTG GCTGTCATTG AGCAACTTGT GGTTCCTGAA
8951 GGAAATTATG TGAATTAGGA TGGTTTTGTA TCATTTATCC TTAAGAACAG

FIGURE 3C

9001 GGAAAATTGA GATGTTTTCT TATGTTTCTG CTGGAGATTT TGGAAAGATG
 9051 TGAACCTAC ACCTACAGAT TGACCTTGCT TAGTTAGCTC TGAACCTCCT
 9101 GCTGCCTCTT CCACGTAAAG TGAAAATTTT GGATTCTTAT CGGCTTCAGA
 9151 TAAACTTACA GGTTAGTGAA ACATAGGGAC TGAGATATAG TAATTCATTC
 9201 TGAAGCTGTT TTGGAGTGGT CAAATAATTT TAGTTGGATA ATATATATTA
 9251 CTGGCTAATG ATTGTGGATA TTGGAAGTGA TGAAAAAATT ATTGAATTAT
 9301 TTCTTTCTGC ATTTCAAATG AAAAGGCTAT TAGTTTGAGC AGAGAATTTT
 9351 GATTTAGTAA ACAAATATAT TAAATTCAT GTTTCATTTT TTTCTCCTAT
 9401 CTGGGTTTCTG ATACTCAGTC TTATAAATGG AACATGATTT ATTTTGTCTC
 9451 CCTAAACTGG TTATTAACCTT CCTGTCCATA ATCACAAAAC TATATAGATT
 9501 ATATATTTCT TGTATTATTT GGATTTTGAA TACTCTCTTT AAAATAATCA
 9551 AGAGAAAATT AGAGCTGTTA GAATGTTAGA ATTTGTTTTG AAGGCCACAC
 9601 ATAGTGTTC CTTCCACAGAG AGGTTATACT AGTAAATGCC TTTCTATTG
 9651 AGGTCAACAA CTATGACAAC TTCCATTGAA CATGAGTTAG TATTTTAAAC
 9701 GTAAAGCAAT TTTTATACCT GTATGCACCC AAAAAGTAAC AGGGGCTCTC
 9751 AAAAGGGGTG GGGATTGTAC TGTTTACATG TATATTGAAG ATTGCTAGCA
 9801 GAATTCTGGG GCCAGCTTGG TGGAGCGGAG TACACTTCAT TGTCTTAGT
 9851 GTAGTAGCTT CCTCCTCTAA TTTTGAGGTG AGAATGCAGA ATCTGTTTTT
 9901 TGTTTGTGTTG CTTTTAGGTA GGAATAAAAG CAAGGCAAGG AATAATTTTG
 9951 ATTACTTGCA ACATTAAACT TGAATCCACA AATCCTTAGG AAGTGAAGTT
 10001 TTTGATTAAAG ATTATTTAAC TGCCACTTTC CTTGAAAGGT TGTTTAAGAA
 10051 CATCATGTAC CTTTGGGTAA CTTCAGTGG TCTTGAATG CAGATTCCAA
 10101 AGTAAGATCA GCGTTGAAGA ATCTTGACCT TTTCAAACAG GTAATTTGTT
 10151 AGTATGTGTA GTCTTCAAAG TTAAGTTTCA GAGAAATTGC TTTGCTCTTT
 10201 TATTCTTTTC CCCAATCAGA ACTGATCTTT ATGTAATACT ATTAAGATCT
 10251 ACTAATTTCC TGAATCCCT TCATAAGCTT AATCTGGCCA GGTCTTAACC
 10301 TTTATGGATT AGAAATTTTA GTACTTCTTA AGCTAGAAGG CCAGGCCAAC
 10351 TAAAGGGAAG CACTCTCCTC TCTGCCAGTT CAGCAACTAG ATCTGTCTC
 10401 AAATACCTGC CACAGGGATA TGCTGCTTGA AGTTGCCAC GCATGTACAC
 10451 ACTGGGACCA AGAAGGCACT TCTGGTGCCA GAAACAACAC TGTGTTTGCT
 10501 TTGTGGAAT TTTTGATATG CTTTAAAAA TGTAGGTGTC TTCTCCCTCA
 10551 CCTCCTGTGT GTAACCTCCA GCATTTCTTT TGTTTGCTTT TTATCTACAG
 10601 AATTCATGTT CTTTGCATTT TGAGTTAGTT GAATCCACTG TATGCTTTCC
 10651 AGATGATAAT TAGTGAAGCT CAATGATTCT ATGCAGTGTC TTTCAGTGGG
 10701 TAAGGGAAGA TATTCATCAC AGGGGTGGGC CTTTCATCACA GGGGAGATGG
 10751 AAGGGCAAGG GAGAGGGATT CTGATGGCCT AACCCTCATT ATTGCCCCC
 10801 ACTGCCACAT GCAGACACTC TTGCCGCTTT CCCATTGATT TTGTGCAATA
 10851 TTTTAAATTT TTCCACCAAT CTTTGAAAA TATTAATTTA TTTTGTAGGA
 10901 GAAAATGTAT TCATTTCCTT AAATATCCCT CCTCAGAGCA ACAAGCATGA
 10951 TTAGTTTTTG GTATACTTTA AAAAATATTT TACCTACTTA TGGGTAAATT
 11001 GCAATAGATC TCTCCCTAT CACCCACTTG TTTTGTACAA ATGAAAGCAT
 11051 ACTATATACA CTGCTTTGTA TGCTGCTGCT TTTTCTTAAC TGTACATGTT
 11101 GAAAGGTTTT ATTTATAGAC CTCTTACNN NNNNNNNNNN NNNNNNNNNN
 11151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 11201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 11251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 11301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 11351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 11401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNC CTCCTTACTT TTTTATGCAC
 11451 TCNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 11501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 11551 AAAAACTTT GAACATATGT TAAAGATGTT GTTTTGCATG TTCTTAGTGT
 11601 TTCTGCAAAA TAAATTTTAA TATGTGGTCA AATGGCAGAT ACATTTGTAA
 11651 TTTTATCAA TTGCCAATG ACTTCTTACT CAGTAGAGAA GATAGTTTAA
 11701 GGGGAAAGAC TGAATGGCGA TGCTGACCCA AAGGCTCCAA CAATCCCAT
 11751 TCCAGAGTCT GATCTTTCCT TGATTTTAGG ATTCTTTGAT CCCTTTTCTT
 11801 TCCCAAGAAA TCCCTCTGAC AACTCAGTGA ATGTCCATCT GCTCCCATCA
 11851 TCTGTTTCCC ATCAGAGCAG AGGCAATGGG CTGAAGCTAG GAGAATGGAG
 11901 TGAGATGTGT CCATTTGCCA GGGTCTCCCA GTGGCTTCCT GCCTATCATT
 11951 TACTTGTAAG AAAGAGCCAC GCATTCCTTT AGGAATTGCT TAATTCATTG

FIGURE 3D

12001 ATCTTTTATT ATCTTTGTAG TTTAGCCATA ANNNNNNNNN NNNNNNNNNN
12051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12201 NNTCAACTGA TGGAGTTGAT ATCCAGAGAC AAGATACAGA GATTCGTAT
12251 CTTGNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNTAGCCATA ATTTTATAC
12401 CTTTGTGTAT AGTATGGGCT TTTTTTTTTT CTAGACTTAT AACCTATTAA
12451 GGAGGAATAT GTAGAATTCT GAGATCTGAA TACCATTTTG GGGACACCGT
12501 TTATGATAAA AAGATTTCCT GGAAATATTA AGACATTTT GTACCAAAT
12551 ACTTTTTGTA GTGTCCTTAA AACCATAGGG GCATATATTT TCTAATAGTT
12601 ATTCTAATCC TTATTCCAAT TATATACTTA AATGTCTCC TTTAGAAAA
12651 GTAAGAAATG CATGATTAGG TGAAAAATTA AAAGACACAT ACTAAGAGCA
12701 AGACAGCAGT GAAGGTAAGT ATTTCTAACT TTCCTGTTTT CCTGACACCT
12751 AGTTNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
12951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNCA TCATCACTGT TGTATACCTT
13101 TAGATATGGA GGACATATCT AAAGACATTT GATTTCTGTA TTAGTATAGT
13151 AATTGAAGAG GTTTCATTTT ATTGCTACAA ATTTCTTAT GTTTAATCTG
13201 TAATGTAAAG GAGAGTAAAT CATGGCTAAT ATTACAGAGT AGAAACCTTC
13251 TGTTCAAGTCA TCTTCTATTT TTGATTCCTT AGAAGATACC TTAAAAGATA
13301 AATTAAATTG ATTTCTTTAT TTCTATTTGT TTATGCCCTA TCTTGTTTGG
13351 GAAAGTATAT AAAGTGGCTT CCAGGAATTC ACACAATTAA AATAAGGAAA
13401 TAGGGGCTAC ATGGAGGAAA ATGGGGGTAA AATAAAAATA TTAGGAGGAG
13451 GGTTTTTCAA ATGCAGATAT GTAAGCCATA GGATCCTACA TATGTCCTAC
13501 AATTGGGCTA GTAGCTTCTG GTAGCCAAAG GTAAAAGATG TTATTGCTGA
13551 TTTAATTTGT GTTGTTTGTG AGGGAAAAAG CAAACAAGTT ATGTGGGAGG
13601 AATTATGGTT TTAAGTCATG TAGTTCTGAA CGAAGACTTT CACGTGGCCC
13651 TTTACGGAGT GGGAGTGAAT GATGGGATGG AAGGGGCCAC AGCAACATTC
13701 CTATCATGAA GCCAGTGGTG GTTTTGTAAA AGCTGTTTCA TAGAAGATCC
13751 GTCAACATAA GATGAGGCCC TTAAGCTAAA AGTACAGTTC ACGGAAAGCT
13801 GTTCTATACA TAGTTAAGA AGCATTATGT GCATAGCTTT CTGACGTTCA
13851 GCTAGATACA GCCCAGAATA TCTAGCTGTG ATGGATGAAC TGTATAGGGT
13901 TTTGGATGGA CAGCTCTGGA TTAGACAAGA TAGTTTCAGG TTAGAATCAC
13951 TAACAATGTT CTGAAGTTTG CTGTATTATT TAACAGATTA AAGGCCAGTT
14001 CATTTTGTCT TTCTTTTTTA AGTCGATATA TTTTGAAGAT CAGTAATCAA
14051 CTAATGGAGT TGATATCCAG AGACATTGTG AATAGAGCAG GGATAGGGCC
14101 CCGCCTAGGC TCCAGTAAAA GGAGGATCTG CAAAGAAATT AGTGGGAGAT
14151 TTTTAAATAT CAATTTTGTA TCCATTGGAT AGCAATATAT AAATTCACCA
14201 CCATTTTGCT ATATACATAT TCATTTCCAT GCCAGTTATG GCTTGAATA
14251 AGGAGGAAAG GCATGAACAT TGTTCACAGG CATCATTTAC CCTACTGATG
14301 CTGATACATG GATATGGGCC AGTGAGCCCA GGAACAACAG AATCCCTGAC
14351 CAGTGCTTAT TTCCCAAAAG TTCTGATTG CTATAGGTGT TCCGGGGAGT
14401 CAAAGTAACC TCAACTGTTT TCTTAATTCA CCTGGATTGA TAAGGGTTGA
14451 TTTATATAAG TTGCTTAGAC CTGAACAGAC TCAAAGCAGA GTCTGTAGGA
14501 AATACTCTGC ATATCAACAT CCCGTACCCA AACCTAAGTC ATCTTTTAC
14551 TGGGGTGTGC GGAAGGGCTG ATTCTCATGT ACTCTGAAG CCCTAGGCAA
14601 TAGAACCTGA AATCCTGATG CACATATACC CTAGGATAAT TTCTCTCTCA
14651 AAAAAAGCA AATAGTGATT TTAATAATTA TGACTCAATG CATATTCTGA
14701 GAGTGCAGCA ATCCAGTAAG TGTCACATCT CCTTGGGAGG AACAAAGAGA
14751 AGGTCTTCAA ACGCTTTTC TGCAGAGACT AGCGTGAACC AAGAATCTCC
14801 CTGTCTGAAT TGTCACCGTA TATCAACGTG GCAGCGAGCT GAGAAAGTTA
14851 GCATTGCCAG GCGCGAGGTT TTCTGTCTTC ACGGTAACAA ACAAATGATC
14901 AGTCCTCAA GATAGTAAAA TGTAAGCCAAG ATAATTGGC TTGAAAAAAA
14951 TTCAAAGTGA TGTCATAAC CAGCTACAAC CACAGAAAAA TACTAGTAGC

FIGURE 3E

15001 ATATTGATGG TGCATCCCTG GGAACATGCC AGGGCAAAT T GGGGTGCAAA
15051 ATACACCAGT TATTTAAAAA TTGGTATGTT ATAAGGTAAA GAATTCATTA
15101 AGTAATATCA AATACAAATA AACTTTTTAT TTTGATTCAC TAAAACTTCT
15151 TTTTAAGTTT TCTGATTTTA TTACTTAATT ACTACTGATA CATAATTAAA
15201 GAATTATGGC CATAATAGAA TTCCTAATAA AATTTCTAAT AAAGCCAGAA
15251 TGAAGGAGTA TATTACAGGC CAGACATGAT AAAGCATTAT GATGTGTGGT
15301 AAAATAGTGA CATCTATTTT TTCATTCTGT ATTATTTTAT AAATTTTCTG
15351 GAGAATTTCA GTTTAAACAG CCTGCTGAAA TACTGTAAAA TCAACCTGTT
15401 ATTCTTAACT CTGATGGGAG GAAACAGATA TGAATAATAA AAATGATTTC
15451 TTAGCTTTAG AATATAGTTG TTGCTTTGGA GAACAAACCG TTTTATTCCA
15501 ATTATTTTTT ATTAGGAGAC TTCATTTTCT GCCATACATT AGCTTTGGTA
15551 GATACTAGAT GCCAATTTG AGAATAAGGA TTTGGCATGG
15601 GTTATTTGCA TGTGAGAGT CAAGATTAA CTAAATTTTA AAACCAAACA
15651 CTTCAATTGA TCAATATCTT TTTAACCATT CTGTAGATTA ATAATATAAA
15701 TTCTCCAGAG CTGACATTAC TTTGCTATAA CATCATCAGA TCACAAGATT
15751 AGGGTTGCCT TTTGTAGATG TTATTCACCT ACTCTGAAAT GTTAGGAAAT
15801 ATGTCATCAC AGTTTAAATT TGTAGTAATA TATACAAAAG GAAAACACAA
15851 CTAGGAATTT TGGATTATG CTTACTTTGC CAAAACCAT GTTGATTTTC
15901 AAAAACCTTT AGCCNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16201 NNNNNNCTAA CCTTATTTTC GTAAATCAAA GTTCCTAAAT TCAGTTTCTA
16251 AGGAAAGCCC CTAAGTGGTC AAAGGTGTGT GTTGCTGAAT TTCAGAAACA
16301 AACTCCATGG ATGGACTTTC CAGACATAAA CTTCCCACCA AAATATTGTT
16351 CCAAAGTGTT TAGTGCTCC TATTTTATTT TAGTGTTAAG ATTTTGTAGG
16401 TACTTGGTAA TTATCAGCAG AATATTTACA TACCAACAAA TAGCAAAGCC
16451 ATTGATAATT AATAGAAAA CAAATCATA GCTTATGTTT AGTTGTATTT
16501 TTTAGTGTAT TTCAGTGTTT AGAGTTGATT TTATATTGCA TATACTAAT
16551 GTGGTAATAG TTTCTACTGG ATAGATTATT TCAAGTTCAC ATGATTATAG
16601 AAGTATTTTA TGAAGTTTGA CATAGATAAT GGTGGTTTCA GGGTCCCTTG
16651 TCTCTTGGGT TGAGTTTTTA CATCATTGAA GAATTAAGG ATTGTGTCTT
16701 GGTGTCTATG TATGGCCAAG AATTAAGACA TTCTGATGTA CAGTTTCATT
16751 TTCTAACTAA GGCAAAAATG ATTTTGTAAA CAATTACTTT TTAGGATAAA
16801 AGTATAAAAT TATGAAGAAA TAATGTGTAG GTTGTCTAAA TGTGTGTTC
16851 TTAAATCTTG TGTTAGGTC TGATGAATTT TTTTATATCT TTAATAGTTC
16901 TAAAGTAGTG TACTGAAAGT TAGGATCATC CATGTTGCTA CTTAACTTGT
16951 TGTCCATAGG GTCTGTTGTC CTGGATGTTG GTGTTATAAA TGTACAGAGT
17001 TAGGTAGTTT TCTGTGAAGA GTTGCAGGC TAATATCACT GTTTTGTGACA
17051 ATGAAAAATA TAGCATGAAA ATTAAGGTTG GGGTATGGAG AACTTCTAG
17101 AGCTATTGCT TTTCTAGCT GATGATTAAG GTTGAGAGGC AGAAGTACAT
17151 TGTGTCAGG AGGACCTTTC CCTTTTATGC ATGCACCTGG CATTATCAC
17201 CTTTAACAAA GTGTGTGTGT CATGCTGTGC TGCTTGCTTA AGGGGCTGTA
17251 TGCCTCCAGT CTGAGTCACA TGGTAACTGC ATCAGCAGTC TTAGCCTGTA
17301 GCATTTTATT ATTTCTTTTC AAAGTTTACA CTTGGCCTAA TACTTAGACA
17351 TTTTATAAAT CTTTACAGTA AATGAACTA TACATGAGAA GATGGGCATA
17401 CCTTTGGATT AAAAAAAAAA AAGGCTCAGT TCTTTAAAGT TTCTTATCCT
17451 TGATTTTCCT AACACGGTCC AAAGTTCAGG ACTGGCTCCA AACCATAAA
17501 CCTGTGTTTA GCAAGCAAGA AACATAATCC CCAGATAGTT TAGTTTCTCT
17551 GGTTTATGAG TCACAATTTT ATAAAATCAT CAGAGTGCTT ATTAATTCCA
17601 ACCACGTATA GTAAAGAACC TTCAGATGAA CTGAAGCAAG GGTCTTAGG
17651 CAAGTTGCCG GAGAGAGTTC TAGAATTCTA AACTACCTGA GTAGCTTGC
17701 TGAAATGTTG CTTGTATTTG CTAAGTGGC CATTTTATGA TGGCCATAGA
17751 GCAACAGATT ATCAAGAGAA AATGAGACAG ATTTTCTCTG ATTATGTGCT
17801 ATGAATGAAT CCTATTTTAG TGAATGTTTT AATGGGGTTT ATACCGCAAA
17851 AAAAAAAAAA TGTATGTAAG GCATTTATTT CTGGCAACTT TCATAAAAT
17901 TGTGTGTGATG GTTGCATATA AAAATTTCTT TATCCTTCAA TAGAGGATAG
17951 TTCCAGAAA CTTCTAGAA GTAATCTATT CCAGATTTAA CATTGCTTTG

FIGURE 3F

21001 AGGACAGATG GTACAGCACA TTTTCAAATG GTTTTGTTTA CCATAATCAT
21051 CTTTTGGAAA TAATCAATTG GAAGTAGAAC AATTCTTCAG TTATTCATTT
21101 CTAATCAATG AACTTTCATG GGATAGTTTT CTGTGCTTCT GAACCGTTTT
21151 TCCCATTTCT CTAATAATAT CCATCCATAA ACCTCCGAGT GGAACAGTTT
21201 TGGTTGTCTA ATACAGTTTG TGCTTGGGAA ATGAATTGAA GGGAGATGGA
21251 ATTGAATTGT TGAATGTGAG CATTTAGAGT TCCTGTAATT TATCCGTNNN
21301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
21351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNTTGT
21401 TAAGATGATT GGCTCTTATT CCTTCTATGA TTCTTAAATA GGAAAAAGTT
21451 AGGGATAAAG ATTTGAGGGA GTTCAGGGAA AGAATCATTT CCTGCTGAAG
21501 AGGATCAAGG AAGACTTCAT GAACTCAATG GTATTCAGTT GATACTATAT
21551 ATATATATAT TCTAGTATAT AGTATCAGTA TTGAAAGCCC
21601 ACAGAGGAAG GCAATACTGG GTAAGAGGCA TGTTTTAGTA TAGGATTCTG
21651 TGAGAGCATA TCTTAAAGGG CAGTAGCCCA GACCAGCAGT GATGGAGTGG
21701 TGATGGCCCA GCTGAATCCC GAGGGATAAA TCAGTCAAGT GAAGAGTATC
21751 CTGGGCACAC AGAAAAATGC ATTTAAAGCT CAGTAGCCAG AGGGAGGAGC
21801 ACATTCAGAG AACTGTAAAG AATCTCAACC CAGCATATGT GGGGTATATG
21851 TTTGAAGGGC ATCTGATGAG AGACGCAGTG GGAGAAGTGG GAAAGGACTC
21901 CTGGAAGGCT TTTGATATCA TATTAATAAC ATTTGCCCTA ATCACAAGT
21951 TGTAGAGGAA CACTGTAAAG ATTTTAAAGCA GAGAAATTAA TATTTTAGAG
22001 GATAGCTGGT GGTAAATGGAT CAGATCTAGC ACTCAGGTGA ATTTAGTTTT
22051 GCTTGGAGTG TTTGTAGTTG TCAGTGTCTC ACAGGAGACA CGGATATAAG
22101 GGTGTTGTTCA AAAATCGGTA GGTCTGCATT GCTGCGAGGA GTCTATCAGC
22151 TGGGGCCTCT TAAGTAGCTG GCCCATTTGG AAGAAGTATG TGCCTCATTT
22201 CCCTGTGGAG CTCACTTCCT TTAGTGCAA TGCCTGGGAC CTGTAGGAAG
22251 TCAACCTGGT GAGAAATTAG GCTTCATTTT AAGTAGTGGT GGTGGGAACA
22301 AAGAGAAGGA GAGGGATTGA AAGATACTGA AAAGAAAGGA TGGAGAGGCA
22351 CACAGTGTTA GGGCAGTGGC GGGGGGGCAC ATGGGCTTGG CAGACAGACT
22401 TGNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
22601 GTTGTGAAAG GTGAAGAAGA TGTAATCAAG ATTACTCCCA AGTCTCTAGG
22651 ACATCATTTG GTGAGCTAGA GAACACAGGA GAAGGTGCAG ATAATGATTT
22701 CAATTTTGGA TATGTTTCTG TTGAAGCGCT TGTGAGACTT TGAAGTAGAG
22751 AAACCTTGATA TGTGCTTGGA TTGCAGGGTG AGATCTGAAC AGAACACAGG
22801 CTATGATTTT ATTTCTTTCC ACTATTCTAC TCAAACCTTC CTTCCTTTTC
22851 CTTCTGCAAC TTCAGTCAAC ATTGATTATT GCCACAATA TCTCCACGGA
22901 GGAACAGCTT TAAACAGTAG AAAAAGAATG TGGGATAGGT AAAAATCTGT
22951 GTTCCACATC TGCTAAATGT TATCTAATAG TTGTAAGTTC TTGAAAAAAA
23001 ATCTTTCCAG TTTACCGTTT TGGCAAAGCA TTGTAAGCAA CTAATAATTC
23051 AACATTTGTT TATATCAAGG GCTGAAAAAA AGTCATCCTC TACAAACATA
23101 TTTTCTTTT TTTCTCTTT TGTTGCTT TTTTGTGTTT ACATCACCAC
23151 CACTCCTAAC TCTAAGAAAA TATTTGACTA AAAGTGAATC ATTGTTAGTA
23201 GTGAATTCGT ACNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23501 NNNNNNNNNN NNNAGCACAT TTATTTTTTA AGTTTATATT CTGTCCTAAT
23551 ACAATGTATT AAAAATGAGC CCTGCCTCAA GTCCTAGCCT AAATGCCACT
23601 TTCTCCAGTC TGCTTTCTCT ATTTTCACAG CAGGAAGTCT TCTGTACTTC
23651 CTCTACATAT AACACTTGCA TGAAGCATAT CTCACATTGC CCTACATTGT
23701 GGTATTATTAT GGATCTATAT TATTTGTTCT GCTAGACTAT TAGTTTCATG
23751 AAGGTGGAGG CAGTGTCTTA TACATCTTTA CGCCAGTATA ACAGCTTCTG
23801 CAATGCCTCG TTCAGTTCAG TTCAGTACAC TTTTGAAGGT GTAAGAGCTA
23851 GTCATTCTTA TTTTATTAATA AAAGGGATT AAAAAGAGTA TGGACAAGAT
23901 GTAATCTTAA TAAACATATT GTGGTTTTCA GGAAACAGTT TAATGAGAAA
23951 ATGAATATAT TTGCTCCCTT ACCTTATGAC CTTATTACAT ATCATAGATC

FIGURE 3H

24001 TTCACTCTTG CTTTAGATTG TAAACAGAAA TCAATTTATG TGCTTGAAAT
 24051 CACATAAAGG TAATCTGATT GTGTCATCTA CTGCTTACAG TCCATTGGCT
 24101 GTTCACAATT ATGGGTAGAA TCTTGATTAT TTGAGTAGAA GCCTTTTCAG
 24151 GCTGACATTG CCTTATCAGT GGAGGACTGT TTCTTGTTGT AGGCCCAGTC
 24201 GCAGTGGACG GCCCTTGGGT TTCCAGACTT CTGCTGCTTC ATGCCTGTTG
 24251 GCTTTTCTTA TGCTGCTCAG TCTGAGTCAA AGGCCCTTCA CCCATTCCCA
 24301 TGGTGAATTT CTACTTATCC TTTAAAATTC AGCTCACATT TCCTCTTGAA
 24351 AGTTTTTCCT GGTATCTTTC CTATTCCTGC CTCCCTGCAG AGAAATGCCT
 24401 TCTTCTGTAT TCTTTTAACA CCTGTACAC CTGTCATGTC TATCTTTGTC
 24451 AGGAGACTGT AACTACAGA GCAAGTACTG TGTATTTAAA GAGCTCAAAG
 24501 TCTAGTGAAG GAGCTAGATA AGAAAAACAA CTCTGAAAAT ATGGTATAAG
 24551 ATATATATAT TTAAGGTAA ACCTGTATAT ACATATAATT TATGATAGAA
 24601 GTATGCACAG AGGCTTATGA AAGCACTGAA GAGGGTATGA ACTCAGCATG
 24651 AGGCGTGAAA ATCTTGGGGG TGATATAATG GTCAAACCTGA CAGGTGAAGC
 24701 CTGTCTAGAT ATTAGTGTAG CAGACAAGAG GGGGAATGGC ATTCCATGTG
 24751 AAAGTGTGGA GGTACTCGAG AGTGTATTTT TAAGAAGTGC AATTAAGTGC
 24801 TATATTTGTG AATGTGATTG GAAATGAGGC CAATGGGATA GGCAAGGGTC
 24851 ATATCATGAA ATATGGTACT GAATGCCATG CTGAGTGTTC GAAATAGTAT
 24901 AATGATCTAC ATGAACATGA TGAGGAGCCA GTGAATGACT ATAAGGAGCA
 24951 GAATGGCATG CTGAAATCAT GGGTTAGGGA GGTCCCTCTA ACAATAGTTC
 25001 AGGCCTGGAT TGAAGTCCAT GATGTTGGAC ACAGAAAGAC TGTNNNNNNN
 25051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 25101 NNNNNNNNNA AGGTTCTTCA ACAGATAAAT TTGGAATAAT TATCATTTTC
 25151 TTTAGAAAAC AAATAAGCAA ACTAAAAACA CCTGAAGGTC AAAATGACAA
 25201 ATCTCTGACA TCATATTAGT CGAGGCCAG TAACAAAATT GNNNNNNNNN
 25251 NNNNNNNNNN NNNNNNTATAC TTCAACTCAT GTATCAGTGA CTCTTTTAAA
 25301 CATAGAGATT TTGGTTGCTA AAAAAGGTGA TGATGGTTAC ATTAAAATTT
 25351 TCTGGCCTTT TGTGATGATT CTTTAGCAAA CCTGCTCAAA ATTACAGTAT
 25401 CAGATGTTTT AAAAAATGAT AATTTTCAAC TGAAAGAAAT ACTGTAATGT
 25451 ATTTTCATCAT ATTCTATGAA TTCTTACATT ATGAAAGATA TCTTTTATGT
 25501 TTCAGATATC TATTATATTT CCTTTTGTAA AGAATTTTTT CTAGCACTAC
 25551 ATCTTTTCCA AACCTTTTGT ACTTTCCTAA GCCTGTGTAA TTAATCACTG
 25601 CCTGTCTGTG GTCCTTAACA TGCATCTGTT ATCAGGACAT CTCTGCTTAC
 25651 TTCTGTTTCT CTTACTAGTA CTTACCTGCT GGAGGATGAG AACTTGCTCT
 25701 CCTATTCAAT TTTATATCCC CAATTTTAA TATAGTGTCT TAGACACTGG
 25751 GGCGTTTCCT CCTGTCTATA CTTATAATCC TCNNNNNNNN NNNNNNNNNN
 25801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 25851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 25901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 25951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNCCTT TCTCTAAATT
 26051 ACATTGAAAA GTTAACAAC ACTCACAAGT TAAATAATAG TTGTCTGTGT
 26101 ATTCACTGTTA TTGTAACCTT TTTTATATAT TTCTTCCATG NNNNNNNNNN
 26151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 26951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN

FIGURE 3I

27001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 27051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 27101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 27151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 27201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 27251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 27301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 27351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNTAC TCTACTNNNN NNNNNNNNNN
 27401 NNNNNNNNNN TTCACATTTC CAAATAAATA GTAACACTTT TTAATGTGTT
 27451 CTTAACGTTT TATTTGGCAT TAATCTAAAT TCCCCTCTAG CACAATGAAA
 27501 ACAGAAAGAG AAAGTTAAAA TTCAAGTAA ACTGAAAACA ACAATAGTGC
 27551 TCAGAGGATT TTTTTTTAAG TGAAAAGGGA TAGTGCTTAA TTATGACAAA
 27601 TAAAAGCTAA CTTGAGATGC ACATATACAG ATGCAGCCAC TTATTTTGGC
 27651 GGGGGACACT TCAGGAATTA AAATTTAAAT AGCGAGAGAT CAAATAGATA
 27701 CTTGGTAAAT GTGTCTGAAT TGGATGTTCC CAGACACAAA AATAAAATGA
 27751 GTTATTGACA GCTCTTGGGA GACAACATTA TAAAGACTAG ACACGTTATT
 27801 TATTTTAACT CTATGTTCTA AATTACCATT GAGTAATTGA CATTTCGTATT
 27851 TGACTATGGT TTGTGGTTAA GTTCTTAAT GCAATAATGT TAAATAAAAT
 27901 GTGAAGCCCA AAGCAAAACA CAACAAAAAT TATAGCAATA CTTCAACAGA
 27951 GGTAATAATA ATATGCTGCA TCAATGGTTC AGAATCCAGC ATCTACATAA
 28001 AACAAGCAAC AGGGTAATGA AATTATTTTC TTTTCAAATA TTCTGGCAGA
 28051 GCTACTTTAG TTTTCTTAAG TTATAGATTG TGGTCTTAAC TGCAACTTTT
 28101 CGCTCCTTTT AAGAAGTATT TAAGTTATTT AAATGTTACT TAAATAC'TTT
 28151 TATGTTTTTA ATCATTATTA ACATTCTCTA CCCCTCTCAT CCTTTCCTGC
 28201 TTAGTATTTT GTTATAATCT CACTCTCCCC ACTTCCAAAT GCACTCAAAA
 28251 ATGCTGGACT TTCTGGTCTT TTTCTGACCA CCAGAAGAAG TAGTGGATTG
 28301 AGCATGGAAC TGGGATAAGA CATTTTTCN NNNNNNNNNN NNNNNNNNNN
 28351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 28401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 28451 NNNNNNNNNN TGAGTTCTGT ATTCATGAGG GCATTTAATG TGTCTTATAG
 28501 TGATAACAAT CAGAGGCATT ATTATTCAGT TTTTACTGCA GCAGGAGTTT
 28551 AGATTTGACC ATCAGATAAT TGACACTATA GTACAGGGTA TTAGGAAAGC
 28601 TGGAAGATG CTTTCCCTCG GAGATCTTGG ACAGTAGGAT TCTTCTAGGT
 28651 TTTGTTCTGC TCATGATATG CAATGGCTCA GGGCACTTTT CAGAGGACCA
 28701 CTGTACAATC TCCCTCCATT CCACCCAGGT CCTATTAATA TAGTAGTGAG
 28751 TTGTTTCATA TTGCCTTCAG TGACTTTCCC TTTTCCAAGG CCAATTACCA
 28801 TTTGAAAAAG TCATTCCCTG TCATATTTTC TTGTTCTGCT CAAATGAATT
 28851 TGTCCATATC TGAATTCTGT GAAGCTTGTT GGGTGTAAAA ACAGCTTTTC
 28901 AAAGTCTTCC AGTTTCATAT CTATTTGTTT TTCTGCCCTA CCCTTGATG
 28951 TTTCTCTGTC CCCTTTTCTT CTGTGAGCAT AACCTGGAGG ACAAGGTTTT
 29001 TTGTTCTGCT TTTAGTGGTG CCATGTCTGC ATAATAACAT GATGTAGATT
 29051 GAAAAAATTA CAAATGATTC TTGGAATTCT AAAGATAATC TTATTTCTAT
 29101 TGAGAAAATC CTTCTCAAGT TACTAACTAC CATTGAGATT GGATTGGCTT
 29151 TTGCCTTATT ACTTTTCAGA GCCTCCATCT GCTTGGGTAT CTCAACATAT
 29201 CCTTAGTTTT CAAATGTTGG CACTTTACTC CCAGCATGAT TACCTTACAT
 29251 AAGAAACATT ATAGGACTGA TGTGGGAGTT TACTTTTCTC ATAACTTATT
 29301 TGATAATTCA CTGCTTATGT TAGAGTTAGA AACTATTGTC CAACTCTCAG
 29351 AGACCCAGTT ACATCACTTA AGATGGATAN NNNNNNNNNN NNNNNNNNNN
 29401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 29451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 29501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 29551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNGGAT
 29601 GGATTTTAA GACTAAAATT TTAGATTGCA ATCAAACAGA AGAGGCTTGG
 29651 ACATGTGACA TTAAAGAGCT ATGTTTGTGTT TGCCAGTTGA AACTTGTTC
 29701 TTTTCTAGT TAAACAATG ATTTAGAGGT TATTTTGAGG GCACTTTGAA
 29751 GATTATGCTA CAGGAATGCT AGAGAGCAAA TTTTGAGAGT GATTGCCCAT
 29801 TTGGACTTAA TTCTGGCAAC TGATTTTGGG GTAAAATGTC TATCATTCGT
 29851 TGATTATCTG TGATTTTAC CTGGACTTTA CTTAAGCTTT ATTAAGTTGC
 29901 TAAACCATAT TTGGATGCTA GTGATAGCAG ACCATCAAAT ACGGCCCAAA
 29951 CTTCTTGTTT TGATCCACAC GAAGGTCAGA GAAGCAATGC TGCCTTTTCT

FIGURE 3J

```

30001 GATAGCCAGT AGCACCAGCA GGCACGTTGT TTA CTCCAAA CAAGAATTTT
30051 AATATTTTTTG AAGACCACTG AAAATGGATC ATTATACTT TTTTATTTTT
30101 TTGATAAGGA AGGGATGACC TACTATTAC AGAGTAATGC AGTTTGCTGA
30151 AAAGGTTGGT TTTTGCTGAC CTCTGAGAGC TCACATTACA GTGGAGTGTG
30201 TTATTGGAAG GTAGGCCAGT TTATACAGAG GTTGGGAATT ATTTTCTCT
30251 AGTTTTGAAA TGGTTATGCA CTTTCTAATC TAGTTATTTA GATGTAGAAA
30301 TAAAGTTTTT TTTTACATT CCCCTGAGTA TATGGTATTG TAGGTATAAA
30351 ATAAATTTGA TGAGTTTTAT TTCTGTACTA GGATCTACTC ATAACTCTTC
30401 TATCCTAAAT TTGTATCAA AGGAAACCTT TGCTGTCTTG ATAATAAACA
30451 CAGACTAAGT CCAGATTCCA CAGTCTAGTG ATCAAGGAAA TTCAGGAATT
30501 GTATTTAGCT ACAAGTAACG TGACACAAAG AACAGTGCCC TAAATTGCTA
30551 GGATGGTGAT TTAGGGTTAA TATTATGACT TCTTGCTCTA CCCTCTTTCC
30601 ATTCTCAAGA TCGCCTCATG GTCATAAAGA GGCCACTGTG GTTGAGTCAT
30651 GACATGCACG TTCTAGGCGA GAACACAGAA CAAGCCGTGC TCTTCAGCCC
30701 CCTTCTTACA CAGCATTCCA AAGCCCCACC CCATCACTTC TGCTTTCATC
30751 CCATTTAGCC AGAACTTAGT CATTTGGCTG CCCATATCTG CTAAGGAGAA
30801 TAGGGAACAT AGTTTTCCAT TACTAACCCA TCCCCACTC NNNNNNNNNN
30851 NNNNNNNNNN NNCCATTCTA TATTAAAGCA GAAGGGAAAG AGATATTGGT
30901 AAGAATCCAG CTGGCCTTTT GTGATCTGTG TCAGCCTTTC TTTTGATCTC
30951 ATCTGCTGTT TTAAGCACTT TACACTGTAG CCCACGAGA ACACTTTGCA
31001 CTCATAAGA AGCAGTCCCC TTTGCTGCGC CCCCCCCCCA CCACTTTGCT
31051 TATATTCTGA AAGTCTTTTG TTTCTATTTC CACTGCTCTT ACCTCTAACA
31101 CACTGCCTCT AACACACCAA CCTGCAGTTG TAGTTATTAC ACACCCTCCT
31151 TGGTCTTTTC ATCTCTCTAT CACAGCCCTT GTTGTGGTTT AGCCAATATA
31201 TTTTAGTTCC ACAGCTAAAT TTTCATACCC TCTATGACTC TCTAATCCCC
31251 TGCCACACTT GCCTACTATA ATACATTATA TATATAACAA ATGTTTGATA
31301 CGTATTTATT GAATTCATT CCAGAACTAA TGCCAGCAAG ATAACTTTGT
31351 GCTATATAGG AGAATATCTT TTTGTGCAAC AGTTTCCAAA GGGTTTTCTT
31401 TTTCTAAGAA GAAAGAAATT GATTGTATCA ACTTTATGAG TATCCTACCG
31451 CATTTAATAG CCATTGGCTA ATCTAAGGGT TCCTGGTTAC TTCCTGAAT
31501 AGCCTATCAG ATGGAAGTGC AAACAACAGT TTGTTTGTAA ATAGGACTCC
31551 CTAAACATGG AAGAAACATT AACAGTGTTG GCCTGTTGGA ATGTGTGCAT
31601 TTGATGTGCT CAAGATTAGG GCACTCTGCT TGAGAACAAA TAACAAAAAA
31651 GGGAGAGGAA ACAATAAAAA CTTTGGTCCT ATAAAGCACC TGAAAGTACT
31701 ATAAATTGAT GGTTCCTAAG CTGGTCAGGG GGTCCAAAGG CTACAGCCTG
31751 GGGGCCTCAA GTTTAATTGT TTTATAAAGT GTCCTAAATA AAATTTTAT
31801 GTTTTAAAGT GGTATTTTAA AAACACTTTT TTCATGCTTT GAGAGAGTTT
31851 TCCAAATTCC AATTATTTTA AGGGGTATT TCTGGACTT GCACTTAACG
31901 ATTTTGAGAT GTTTACATT TTTTCAATAT GGCATTCTGT GTGCCTCAGT
31951 GATACATGGT TATCCAGGTT GCATGCATAT ATAAATGTTA AGATTTATGG
32001 AAGGTCATCT TTTTAGATTA AAAAGAATTT TTTAAGCTG GTATTTCTTG
32051 GTGATAGGGC CTAGAAATTA TGTAGAGTGG CTTACTTCTG GAACCTATTT
32101 TAACTGCAT ATAAACCATC CGCCTAGTGT ACAGTTGGCT AAAGAGTAAT
32151 ATTAGAAGGC CCTCTGGAC AGTTTATTTC ATTTCATGGA TATGAACACA
32201 ATGTTTCCC TTTGAATTTA ATGCCATGTT TAAAATCAGA TTTTAAGAAT
32251 TTTCCAAGGG CATTTCCCTA TCATTTACAC TCTGCTTG TGTTTTCTTCT
32301 GTAGTCTTTT ACATTAAATA CCTCTACAG AGCACTGCCT AAGGATTTGT
32351 GGTGGTACAG GGTCCAGTTG GGATGACAAA CAGGCAAGGA AGGCCTGGAA
32401 GTAAATTAG CAAAGAGGCC CTGTGGAATG GAAGGTGAGG GAAGGGTTAG
32451 TGACAGTTGG GGGAGGAAAG GTAGAAAAAA AAATAACATG CACATCAGTT
32501 TCGCAGGAGT ATTAGAGTCT TAAAGGAAAC AATGTTTGAT AATTATCAGA
32551 GAGGAAACTG GGAGCATAGA GCATATCCCA GAATGGAGAG CAGCATGGGT
32601 TAAAATGGGT AGGAACAGGT GTCAGGAGCT TCAAANNNNN NNNNNNNNNN
32651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
32701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
32751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
32801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
32851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
32901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
32951 NNNNNNNNNN NTGATTTTTA TCAGGCAGTT ATGATGATAA ATGTATGGAA

```

FIGURE 3K

33001 ACTTCCCATT TCCTAGAGCT AAAGTGCATG TTTCTCATTC TGAAATGTAG
33051 GGAACATAAT CATCTGATAC CACTCACCTG ATTGTTTTCTC ACTCTTCCTC
33101 CACCATTTAC CCATCTCTTT AGCTTAATGA GTCCCCTGTG TATCTCCCAA
33151 CTAAACAGCG GCTTACTTGC CTGTGAAATA TTCTTCTCTT GGGTAGTCTG
33201 CTCCCTTCTC TGCTTACTCA TGCTTCAAGA TTCAACATAA GCCTCCTCTA
33251 TGAGGCTTTC TGCACGTATG TATATGGATT TGCTTGTGTA ATGATTTCCT
33301 CACAGATTTC ATATTGCTGA TAAATAAATA TTGTTTTGAA TAAGAAACGT
33351 GGTTTTGTAT TTTTATCTCG ATTGTAGACT CCTTGAGACC AGTACCATGC
33401 TATACAATTA TTTTTCATCT ATTATAGTGT CTGGCATAGG GACATGCACA
33451 TATTTGGTAC AGAANNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
33501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
33551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
33601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
33651 NNNNNNNNNN NNNNNNNNNN NNNNNNTAGAG AATAAACGGA ATAAATTCCA
33701 ATTAAACAGT GAAAAATATT GAGTGATGAT TTACAGAAAT TTAAAGTCTT
33751 GAGAAAGTGG AACTCAGTTG ATGGAAAAGA GTGAATGTCA AACTGAGAAC
33801 GTCCTATTTG CTATGTTAGG GCATAGAAAG GCCATTATAG ATTGAAAAGC
33851 AGTTGTAATG AATCTTAGGG TTAAGAAAGA TCAAGATACA GAAATATCCT
33901 TGAATTGAAG TTCCAAAACA ATGTTGTTTT GGTTTTTGTT TCATTTTGAA
33951 TCCTTTCATA CTTAGGAATA CCATTTCTAG TAAATAAAT ATTTTATGTT
34001 TAGTTAGAAA TTTATCTGTA TTTCATACAT ACTTAGTACT TTTGGACTAA
34051 GAACTGCTAT TTGAAGTATA TTTGAGGGAA TGAGTTTGAA ATTTTGGGTG
34101 CAGGACATTA TAAAGTTGTA ACTATGAATA AATTTCAAGT ATGCTTATGC
34151 ATAGTTTAC CTAGTTTAT TTGTCTATTT GAGTATTGTC CTTGAATTTA
34201 AAATTTTTTT CAGCCCCAAC TGATACACAC ACATATACAT ACATAATACA
34251 TGTGTGTGTG TGTAGCTTAC AGAGTGTTTA TAGGAACTG ATTTTGTATA
34301 CTTTGGCTAC TTTGTTGTAA GTTCTAGTTT TTTTCTTTT ATTATTAAAC
34351 TAGTGCACGA CATCAATGCT ATATGATTGG TGTTTCGTTG ACCTAGAAAT
34401 AATGCATGCC ATCTTCTTTT CACAGCTGTG TGCCAACCAC GATGCAAAACA
34451 TGGTGAATGT ATCGGGCCAA ACAAGTGCAA GTGTCATCCT GGTTATGCTG
34501 GAAAAACCTG TAATCAAGGT AGGAAAACAG TCTGACATAA ATACACAATC
34551 GAAGACACCT CTATCACTCC CAAATTAAAA ATATTCTTAT CTCAAACCTAC
34601 TTTCCATGGC TATTTTTCCA AAATATGTGA GCTGCTATTT TGCTGNNNNN
34651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNTTTA TCAAAGGACA
34701 ATTAAACAAA TTGTATCCTC TTATTCTCTG ATACTAATTA AAATGTATTT
34751 TGAAGAAAAG AATCCCTGCA TCAGTAATTT AGAAGCCTTC TGGTACTCCT
34801 TGTTTTTCCA CTAAGTAGTT GGAAATCCTT GGCCAAATAA TTAACCACCT
34851 CAGACCCAG GTACTGCTTC CCTTTAATGC CAAAGTCAAG TAGGGGATTT
34901 GATTTGAAAT TTGGAGTTTC CTCCTAACTC TGAGCCCTTC GATTCTGTAT
34951 TAAATCTCCC TTCAACTACT GACCAGTTTG GAATGTTTCC ATGATAAATA
35001 AAAATGATTA ATTTAGCAAG CACTTTTTAA AAAATCAGC ATCAGTTGTT
35051 TAAAGCAAAT ATTTATTCAA CACTCATAGT GCCCCAGTA TTTTATATAT
35101 ATAATCCATT TTTATCTTTA CAACAACATT AGGAAGTAGA GCATTATTTA
35151 TTATAATNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
35201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
35251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
35301 NNNNNNNNNN NNNNNNTTTT GGATAAAAAA ATATATTTGG AAATGTGATC
35351 CTCGAACTCA TGCTACAAAG TCAGACAAGG CTTGCTCTGT TAATTAAATT
35401 CAGTTAAAAA TTTCCATTAT TNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
35451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
35501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
35551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
35601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
35651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
35701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NCATTATTAG GAATTAAACT
35751 AAAGAAAGAT TAGGGTTAGA TTCTACTAAA TAGATAAGTG TCAAAAATAA
35801 AAAAGATAAG TCAATTTTTA CTTATTTTTT AAATTATACT CCCTATCATC
35851 TTAAATGTCA GGTGAAATAA TCATGGTGTG TAACTATCCC TTACATACTT
35901 ACTTGACCTC ATCTGATATA GAAATGATAT TGCTGAAATA CTACTGTTCT
35951 TCAGTGTCTG ATACTTATTC CAAGATACTC CTTGAGGTAT GTCATGTAAA

FIGURE 3L

```

36001 CTAATGATTT ATAGACACAG TTTT TTTTTC ACTATTTATN NNNNNNNNNN
36051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
36101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
36151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
36201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
36251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
36301 NNNNNNNNNN NNNNGACACC GTTTTACAA TCATCCAGAA CATTGATTTT
36351 TGAAAACATA CTCAAAATGG TGGTTCATGA ACCATCAGAT GAAATTATTC
36401 ATGAACCATC AGATGAAATC CACACTAGCA AAAATGACAT GTCCCAATGG
36451 TAGAAGCTGC TATTCTAAAAG TGTGTTCTTG TATTATCTGC TATTTGATTT
36501 AGAACTGATT TTCCCAATGC TTGTTTCTC TCTAATACTC TTTTAACTT
36551 GAAATTTACC CAACATACCT GCATAATCTT TTAAATAAG TGATGCTCTT
36601 ATTATCTCAG TTTGCTCCTT AAAAAACTCC ACTTGATTTT TCTTCCCAG
36651 CATAAGTTTG CAGGTAGCAG TGTCTGGTT ATTGGATGCC AATGTTTCATC
36701 TTAGAATCTC AGTATACTTT TTTTAAAGTG GTGAACATAA GCCCTCAGAG
36751 GTTTCATTAA TATATCAGCA CGGTAAAATA TTGTTGTCCA AATGTGAGGA
36801 TATAAAATAT AAAGAAGACG ATTAAATATA GTCACTTTTC AGATTGCAGA
36851 GAAATTCATT TAAATTTCTT TGCAGAAGCT GCATCAACTT TAACATGCTT
36901 TAAAGACATG CTAAGCATAT TATTGAAACA AAAGTTCATC ACTGGGTAGC
36951 TCCTATTTTC AGAAAACCAG TTTAACATGT ACTTTTTTTT CCTGTAACAC
37001 ATACCTATTT CTCTAAAGAA AAAATCGAAT GTATCAAGTT AAGATCTTGC
37051 TCCCCAAAAC CATATTCCTT AAAGAGAAGA GGATTTGATT AGTAGCAAAA
37101 TGGGATTTTA CACCTGCAAA AATAGTGCTG ACACTGAAAA TGTAATTTCA
37151 AGTCCAGTTT TGAAAATAAA ACTACAGCTT AAAATAAACT TAGGTGTTTT
37201 AATCATCTCT TATTTTGCCC TCCTCACAAA AAAGCAGTGT GGCAGGTTCC
37251 TGATGGCAAG GTTNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NAGTTTCATC TAATGGTTCT
37401 TAGTCCTGAA CACTGTGGA GCTTGAAAGG GTTCCACCAG TTCTTGGTAC
37451 AGCATCTTAG AGAGTCTCTG CATCCATAGA CTTGAGGGCC CTGTCAGGAT
37501 CAGGGAAGCT GCCATGTGTG GCCTGGTTGA GTGTGAGAGC TGCCTAGAGA
37551 CTTCCATAAA AGTTGTTAGA GAAAATTGTT GATAGTGCCT ACATAGCAGA
37601 TTAACCTAAA CTGTTTTTCAT TCAGGCCAC ACATTTTAAT AAAGTAGAAA
37651 ATATGCTTCA CAGATAAGGG AAATCAAACA GGCTCCTTTT TTCTGGAGGA
37701 GAGAAATGTC AAAAAGAATT AAATTGAAA TAACTTTACA GAACTGGAAA
37751 TTAGCTTTTG ATTAAAAGTA GCTTTTGGTA TATGACAGGT ATTCACTGAG
37801 AATTTTGTAG CGAGTTATAT ACTTAAAGAA ATAACCCCA GAACTTGCA
37851 TCATGGTGTA AACAGCTTGA ATAAACAAGT GCTTAACCAG TGCCTTTAGA
37901 GCTGCCCTGGG AATACAGCCAG AATACCAGG CAAGCTGCAT TTTGGAAC TG
37951 GTTTAATTTA GTAGCCTTGC CACAGGCTTA GTGTGATCTG CTTTGGTGG
38001 CTTGATCTTC CCCACTAAGT CATTTTCTGG ATTTGTTACA CCTAGAACTG
38051 TTAGGAAATT ACAGGCTTGG GCTGATCATT AACATACTGT ACTCTACAAG
38101 GCACACGTTA CCTTTCAAAG CAGATGAAAT TCTAACCTGA ATTCTGGCAA
38151 GATTCTTTGA TCATTGCTT CCTTTACTTT TACTTTTATT TATGCATATT
38201 TCCCTCCTCC TTGAGTTTCT GTACCAACAC AAACCTCTT TCCCCCAGC
38251 CAGNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NTTTTCCATT
38301 GGGGAATGAT TTTTAGTATG TAAATATATC ACTGCATATT CTCAGAAATG
38351 AAAGACATTC TTAGGAATTT ACAGTGACT TTATAATAAT TTCAGAAGAA
38401 AATATTTATA AATGTCAATT CCTAATGTTT TAGCATGGTT TATGTTTCAT
38451 ATGTTGAATT CTTTATCATA AGGAAAGAAT TGGAGTCTTT TAGGTCAGAA
38501 CCAGATACTA ATTTGTGGA CTAGTTACAT CTGAAAGTTG ACTGCTTTGC
38551 TAAGCACAAA AATCTAAGGG CTTTAACTCT AATATTAAGG TGTTTACCTA
38601 CAGCCGTAGG TTTTGAAAGA TGTATGGTTC CNNNNNNNN NNNNNNNNNN
38651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
38701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
38751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
38801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
38851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
38901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNGAAG TTGTATGCTT
38951 TCAACTCCCA AAACCAATGT TCATTTGATT TGGACATTTA ACTGGCTATT

```

FIGURE 3M

39001 AATAACATAA TGGTGTGAA ACAGAACGTG ATTGTTAAGT CTCAACTCTC
39051 TTCTTTGAGT CTCACCCATC TTATATTTTG GTAAGCTAAT AATGGCAACA
39101 GTTAACATTT TTTTGCATAA AGCTTCTCAT GTATGTTCCA GAAGCATAAT
39151 TTCTTACATT TAAATGATAA TATTGACGAT CTATTATGCT GTTATTTAAA
39201 TAAAATCTTT ATGTTAATCA TACATTTGTG CATAATTTTG GAGTATTTTA
39251 AACTACTATT GTACAACGTG ATAGTGTAGG GAAAAAATC TTAAGTTGGT
39301 CGTTTACAAT TCAGTGTCTT TAAATAGAGC AGGTGGATTA TAATCAATAT
39351 TTGTATAATT TGTTTTTTGT TTTCTTCCTT CTGTGTATGA CTAAATAATA
39401 CTTTTAAAAA TGATGATCTC TTGTTTGGAA TTTTTTAAAA AATTGTATGA
39451 CGATGGATTG TTCATTTATA ATGTGTGTGT TGTCCTTGGG AAACAAGATG
39501 ACAGTGATAC ATTTTTAAGA AAATTAAAGG AAAAGAGAAA CCAGAAGAAC
39551 CAGAAACTAT ATGTTAATAA GTATCTAATG CATTCTTTTT AACTCTACC
39601 AACTGTATGT AATTGTTTAA TATATACTTG CAAAGATTTT TAGGGCTAAA
39651 ATTGACATCA GTTCAAAGTT GACTCTTAT ACCTTCTTCC TGGTGTGAAG
39701 ACGAGCACAT CCCAGCTCCT CTGACCAAG GCAGTGAACA GCCTCTTTTC
39751 CAACCCCTGG ATCACCAAGC CACAAGTTTG CCTTCAAGGG GTGAGCGTGC
39801 ATTGTCCAG GTGGTCTGCT CTTCCACCT GCATGGCTTC CTGCAGTTAT
39851 GTCACCTCTT GCAAAATGTT TGAAGACTAG CCCAGGAGTG TCCAGCCTTC
39901 AGCCCTGAGC ATGAGCAAGC AGCAGGAAT GCCACCTGGT GCTGATTCTC
39951 TGTGCTCTGA AAAGCATGGA GCCCTACTGT CCTTTTAATG CATCGAGTAT
40001 CATTGCTTTT CTTTCTCTTT TCTGAATTTT GAAATATGTT CTAAAAACAG
40051 ATAGCAAAC AACTTTTTTC ATATGTAGGC AAATTCAGGA GTTCCAAAGC
40101 AACTGAAATT GGGGGTTTGA TGCCCTTTTG CAAGGAGATT TTAGAAATCA
40151 AGTCCTCTTC CCCACTCTTG AGGTATAGAG TTTTAGGTTA TCTGAAAAATC
40201 AAAGCATACA TTAGTGACCT TCTTTTGGTA TACATTGTGT AACATCAGAT
40251 AGCCTAAGTT CTGGTTCCAA GAGAGCACTT AGTTGCTATT TGATTATAAG
40301 CAAGTCATTA ACCAATCAGT AACAGGGAGA CAGTAATTCC TGCCCTATTT
40351 GCCTGTGGGG TTGGCGTGAG AATAGAAACA CCAAAAATAT TAATGTCTTC
40401 GTCATGCTGG AAAGCTTTGT AGGAAAATAA TGTGGTAGCA TAATTTTACA
40451 GTTCCTCTTT TAGGTCATTT TATCTCTAAC CCATTCATAG GTTTAAGAAC
40501 TTAGAATAAC TGAATTAATA TAGTTGAAAT TATTAATCAT TGTCTCTGCT
40551 GAGAAGAAAT TTGATGCATT TGTATATTTT CTACACAAGG AATTAGGGCA
40601 AAAGAATTAA TTTCCGGTCN NNNNNNNNNN NNNNNNNNNN NNNNNNGTAGT
40651 TGACATATTG ACATCTGTTT GCCTATGACA TTGCCCGATT TAAGCACCAA
40701 AGCGGGGAGA AAGTCCAAAT GTGTTAAAAC AAATGAAAGT TTATTTTAA
40751 AAAATCCAGT AAGTTAGTTA TCTCCCGATT TTTCAAGCTA CTTTTCAGTG
40801 TCTGCATGCT AATAAAATTT CTGATTTTTT TCCTGAAGTT TAATAATAAT
40851 GCTGTGCAGC ATTGCTTCC ATTCTGGACC AGTTCCCTTC TCCCTCTCCT
40901 TCTCCCATCC AAGTCCTGA ATCCACCAAC CTACTGAAAT GTATTCCTGA
40951 CAATAGTAGA TGCACTTTAA GACTTGTATA CATAATAACT GAAGCATTTG
41001 AATGTAAGTG GTTTATTGTA AAGTTCTATC CTTTAGTGTA AGATAAGCAC
41051 ATGAAATAAT TCATTAAATT TTTTATTTTC CTATTTTATA GATTTTCCTT
41101 ATATATATTA TAAACCTCCA GAGAAAAAGG AAGATAAGTA AATTTAAAAAT
41151 AAAAACACCA AAGTTTATTT TCTAGTTTCT TTTATCAACT TTTAAGATTT
41201 ATTTGAGACA GTATGATCAA TGACTTCATT TTGTTCTGCT TATTATGTA
41251 GGAGTATTTA CTATAATTTG GAAGTAATTT ATTTTGAAT TTATTGCTTA
41301 ATTGAATGAT CTCCAATAGA TTGTGATAAT GAACACAGCA TTTATAGAAA
41351 GCAGCACATA TTAACCTACT TAATATGGCA CTAGGTCAAT GAGAAAAGAA
41401 GGTAACATAA TTGAAGACAA GAAACTCTTA AGAAAAGTGA GGACAAAAAG
41451 GCTTCTCACC AGGACACCAG ATGCATTTAA TCTTTTGAAG CTCTGTACTT
41501 TAGGAAAAGT CTGATATTTG GCAAATTTTG ATAAACATGG ATGACTATGG
41551 AATCCTATTT TGGACTATCT GAAGTGGCTT TCATAAGGGT CATTGTGAAG
41601 TTTTtaggag ACACCTGCCT GTGGCAGATG GGACAATGAT GGCAGTCACT
41651 AGTGATATTA ACACCAGTCA GCTGTCAGGG AATATCATCC AGACCATCAG
41701 CAGCTGGTAG AGTACAGCTT TCTCAATTGC TTTCCATGTT TTGGTACTT
41751 ATATGCCCGT TAATAACAGG TAAATAGCC AGTACATCAT TTCCACATTT
41801 ACCCATTGAA TGTGTCATGT TTTCTTCCTT TCACATATTC ATACAGTCCA
41851 GATTTTTTTT TGGACTCATG ACAGCACATT GGCTTTTCTT TCCTTTCAGT
41901 TTCATGATTC TTAACCCCAA AGTGCTTTTG CCATGGGAAC GGAAGGATAA
41951 ATTTCCGTCG AAGCCATTCT CTAAAACCAC CAGCAGCTCC AGTCCAGATT

FIGURE 3N

42001 ATGAATACTA CCATAGCAAT CCTAAGGGGC CCTTTATTGG CCCAGACTGG
42051 AGCCAAATTG TAGAAGAGCT GCCCAACTG GGCTTCAGGT TTAGTTCAGG
42101 GACATAGTGT CTGAAGAAAT TTCCATCAGC ATAAATACTC CTTGTTTATG
42151 AGCTGCCTGA AACTGTAAAC ACCGAATCCA TTCCCATCAG GAACTCACAA
42201 AAGTTTCTGT TATGCTTTGA AAATAAACCT AGGGATACTT AACTGACATG
42251 TAAAGAAAAA TCCATCAGTA TCCTTTTCTC GAGATATAGG TTTTGATTTT
42301 ACTGTGTTAT GTTGTGGTTT TGTCTGTGTT TTGTTTAGTG CAGAAAAATA
42351 CTTAAAACAC AAAGCCTTTC TCTAACACCA TATTTGGTTT TAAATGCTAT
42401 TTGCTATAAT ATCAAGAAGA TTTATCAAAG ACAGGTGTGA CTCTGCAGGA
42451 CCATATTAGG AACAGTCATT AACTCCTCTT AGAAGATGAG AAAAGTTTCC
42501 CTCTCATTTT TAGTAAATTC TGAAATACAA AGTAGAAGAA TATGGGAAGG
42551 GTAGAAAAAT TCTCAGTTGA TCTTCCTTGA TTTTGGCTTC AATTAACAGA
42601 GGGAGAAAAA ATTTTGTGATA CTGTAATCTG CATGGCGCTT TCTGTGAGGA
42651 AATTTTTTTT GAAATATTTA AAATGTGATT TTGTTTTAAA AACCCTTAAC
42701 TAGTCAAGAG ACAATAAAGT AGAAAATTGA AGAACGCATA GGAGATGACA
42751 AAAGTATTTG GTTTATTTAT TTAACCTTTC CGGGATTTGC CCAGCTCGGG
42801 GTCTGCTTAC AGTTGCATTC AGGGCAGATA AAAGACCTAC TTGGAAAAATC
42851 AGTAAGATAT GAAAAATTTT TTAGATGTAA TAATTATTTG GTGGTTTAA
42901 AACATAATGG AACTTGATGA TTTAGTTAAA TGAAATATAC AGATTTCTAA
42951 TAAATGAGCA GATCAGATTG AAATAGATTA GATTTGAAAA CAATTTTGT
43001 CACTAGACAT ATTATCTATA TTTTATTTC AATGACATGG ATTAATAGAT
43051 AATTAATTA TGCTAAATGG GTACTTATGC TGGTGGTTCC AGACAAACAT
43101 GTGGTCACAT TTCTTTTAA TGAAGCATAT ATTTTGGGCA TTACTATTTT
43151 AATGTTTATT TGTCATTATG AGTTTATGT TCACTTGAGG CAGTGTTTAA
43201 TTTAGGTTAA AACCTTTTCA TAATGTAAAA TTTGTTGATT CATTTCATAT
43251 TGATACTTAA ACCAATAAAC CTGAAATACA TGACAGTCTA TTACATTTTG
43301 TCCATTAGTG TTTCCATATG CCTTTTTTAA CTGCGGAATA AATGAAAAATA
43351 TGTTGGTAGG TTTTTTAAAG TTCATTTTGT GACTTGAATG CATTACTAAA
43401 GTAGGAACT GAAGTTTCTT CTTAATCTG AAATCATATA AAATCTACCA
43451 GATTGAACAA GAAAACTAC AATATTGATC ACTCTTAATT TATTTCGTTT
43501 TTATCCTCTG AAAATATAGC CTATGTAGCC TTCCTTTGGG ATGGGAAGGT
43551 CAAAGCAGTC AAATGTTTAA AAATCTGTAT ACTTCTCTGA AAATAAAATA
43601 TTCTCAAAGT CCAAGCCTTA AGGAGGCCAA TGCCTTATCC ATTTAAAGTA
43651 AAATATTCTC ATGTAATATG TTCTTTAAAC AGAGAAAGGA AGACATTGAA
43701 CGAATAGGGT TACATTGTCA GGATTGGATC TAGCAATAGA ACCCAAATAT
43751 TTTGAGAATA TTGGCAAACA GTTGTTAGCT GATAGAGCAT CTCTGCCCAT
43801 GAGGACGTTA GTATGCTGCT GTTTCTAGGG TTGATAAGGT AGGCACCTGT
43851 TGTTGAGAAT TTTCTAATC GGCCTTTATA AACAGAGATC CATAAAGGTC
43901 ATGGCGACTC CTGGTTGCC CTAGTGTAAC GTACCTTCAA ATTAATATTG
43951 CATGAACAGG ATCTATCTCT TATGTAAATG ACAGTTACTG TAGCTTGACC
44001 TATTTTTCTC TTTTATTTT TAATATAGAA AAGAATTTTA GATGGCTTAC
44051 ATTAGTGCTG AGGTTTGTG ATTTTCTGCT TTCAACCCAC ATTGTTTCA
44101 GATCTAAATG AGTGTGGCCT GAAGCCCCGG CCCTGTAAGC ACAGGTGCAT
44151 GAACACTTAC GGCAGCTACA AGTGCTACTG TCTCAACGGA TATATGCTCA
44201 TGCCGGATGG TTCTGCTCA AGTATGTCAA GAATCTTAA TGTTTTATAA
44251 GTGCTTTGGG CTTGTTTCTG TTGTGCTCTG AGAGCTTGCT TTTGTGAAAA
44301 TGGCCTCCGG GGTCTCCTA AACAAGATGT GTGCGTGTGT TTATACTTTT
44351 GCCTGAGGAA TTGAAATCA AAATAAGAGG CACCATTTTC GATGTATAAT
44401 ACTCTCTAT CACGGGTGCC AATATTAAAT TGATTAGGAA ATGGTGTTTA
44451 GAATAAGAAT TACAAAATA AGTATGTTTT CTAATTTTTT TTTTAAACAT
44501 GTTTACCCAT TCTTCTATC TGGGCATAGA AGAATGATTC CAGGCTAAAA
44551 AAAAAAAAAA AAGAACACTA CTGACTATTT AGGGACTGCT CAGACAGAAT
44601 CCATCAGGTC TGTTTGTAA TTTTAAATCA TTCATAAATA TTTTCTTCAT
44651 TTTTATATTT CATTATAAAA GCCTTTAGGC TTTTATAGAA TTTTAGACTC
44701 TAATAATAGT CTAAGACTTC TAAAAACAA TTTCAAATAA AAAATAAATG
44751 AATATGATAC CAGAAAAGTA GGATTCATGT TTATAAGGAA GAAATGCATG
44801 TTTTCACCAT CATTGGATG TAAAAAATGG ACTTTGCCCT AAAATTCTCT
44851 ATAGGCTGAC TCCTGAATTG TGGTATTCAT ATGATATCTG ACAATTAATG
44901 ACTTGATTTT TATTTTACT GAATTATCAT TTAAGTGTG TGCATAAAAT
44951 AAAGTTAGCT AATTTTATTT AACATTGCAT AAGTATAGCA TGGAACTTTT

FIGURE 30

45001 TGG AATTAGG TGAATACATG TTTAACATTG TGCAACTCAA TGGGAAATCT
45051 GCTGTTCCCC TAGAGAAATT TCATGGGCAT TTGAGACAGT TACTTGATG
45101 ATTAGTTAAG AATAGCGTTT AGCTTGAGTA ATCTGGAAAA ATACCTCAAT
45151 TCTTCACTTT CCTTGGCCAC TGGAAAAATT TCCAGATAAT CTATATATGA
45201 TAGTTTATA TTTGTCTTTT AATGCTTTCT TTCCTTCCCC CTTTATTTT
45251 CTTTAGATGT TTGATAGTTA AGGCCTCTTT GTCTTCATTG TTGTCAGTG
45301 TCACTGATTT CAGGACAAGT GAAAAACAT AGCTATTTCC ATTACAGATT
45351 TTGTTTCTA CCTGTCTGAG TCAGCCAGTC ACTCTCTTTC CAGGTGCCCT
45401 GACCTGCTCC ATGGCAAAC TGCAGTATGG CTGTGATGTT GTTAAAGGAC
45451 AAATACGGTG CCAGTGCCCA TCCCCTGGCC TGCAGCTGGC TCCTGATGGG
45501 AGGACCTGTG TAGGTGAGTT GTAAATCAA GCATCTCTGT CAGCAGCCTC
45551 TGTAGGATAA AGGGAGAAAG TGAAAGGTGA TGGGAATAAG GAAAAAAG
45601 GCAATTACTT ACATCAGATA ATTAGCTATC GTTCAGAAGA TATCAGATGT
45651 CTCAGAAGAG GACATCTCTG TGAATGGATA ATGGGAGCGT TGTTTTTTAA
45701 AAAATGAGAA ATANNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
45751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
45801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
45851 NNNNNNNNNN TGTATGTGT GTTCTTCAGG CCTCAGATTT TTTCTGGTA
45901 AATTAGGAAG GCAGTGTAGG ATTGAGAGTC TGGANNNNNN NNNNNNNNNN
45951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNA AGTGTTCGGT GATTAGAGTN
46151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
46301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNGTCTATT
46351 AGCCTTAAGG TAAGACCTGT ATCATTAGCC ATGTGAAGAA ACAGCCAAGC
46401 AGCTGAAGAC ATTGCAGGGC ACTGCCGTGA GAGAGAGGTT GGCCTGTGTG
46451 AGGTTTTTCA AGGAAGAAGG TCCATGTAGC TCAAGTACAG GGTGTGAGTA
46501 TGAGGGGTGG TGGCTGGAGA TATAATGGGT GAAGTAGGTG AGGGTGATTT
46551 GTTGCAAGGC CTTCAAGGAC ACAGGGGAGG AATGTGGATT TAATTCAAAG
46601 TGCAAGTAGA AATAGTTAAA AGATTTTGAG CACAGTAATG ATCTGATTGT
46651 ATGTTTTTAAAT GGAATCATTT TGACAGTTGG GTTGATATG CTTTGGAGGG
46701 GGTAAGTAAC AGCAAAAGGA ACTGTTTTGA GGCTTTTGTA ACAGTCTAAC
46751 CCAAGAGATG GTGGCTTCAA ATAGGGTAGT GGTGGTGGAT TAAGAGAAAA
46801 GTAGACAGAT TCAAATACTT TTTTTFAGAG GTAGAATTAA CTGGACTTGA
46851 TGATATGGGA GTTAAAGGAA AGAGGGTTGT TTCTAATTCA AGAATGACTC
46901 ATGTTTCTAG TTTAAGCAAC TTGGTAGATA CAGGTGCAGT ATGTGCTAAT
46951 ATATAGAGTG CAAGAAAATA ATAACCTGGC TTAGAAGGCT GGGTTGAAAT
47001 TTGGATATGT GGAATTTATG TTGCCTGTAG GGACAACCAG CTAGATACCG
47051 TTTATGATGA GCCACTATAT GCTAGGCAAC CATTTGTAGT CAGTTATTTA
47101 TTTTGAAACC CACTCTGTCT TAACATACCC TTGCTCTTCC TAAATGCTA
47151 TTGACTTATG TTTCTAGATG TTGATGAATG TGCTACAGGA AGAGCCTCCT
47201 GCCCTAGATT TAGGCAATGT GTCAACACTT TTGGGAGCTA CATCTGCAAG
47251 TGTCAATAAG GCTTCGATCT CATGTATATT GGAGGCAAT ATCAATGTCA
47301 TGGTAATGAA ACCCAACCAT TGCTTTGTGT TGTTTCTTCC TAGAGCACTG
47351 AAAGGTCTCG TAATTGTGGT GATGGCTGGA ATGTCAGGGG CAGGGGAGAG
47401 TACTGGCGTT AAGTTAAACC AACAGACATC CAGTTTAAAC ACTGGTAGTT
47451 CTCAGTCTAC ATGTAGTTTA TTTCTTCTGT TTATCTGCCA ATTTTATGTA
47501 GATCATCACA TTGCCAAAAA AAATCATTTT TGAACTGTA TATATTTTTT
47551 ATGTCATCAT ATTTATCTCC TAAATAAGTC TCTTCTTTC CACTTTCTG
47601 ATGCAGACAT AGACGAATGC TCACTTGGTC AGTATCAGTG CAGCAGCTTT
47651 GCTCGATGTT ATAACGTACG TGGGTCTTAC AAGTGCAAAT GTAAAGAAGG
47701 ATACCAGGGT GATGGACTGA CTTGTGTGTG TGAGTAGCAC TTGTCTCTCA
47751 GCTTTAAATT CTAGCAGGAA ATACAGGATT ACACAAAGGC CATTGCTAGG
47801 GAAAAAAGG AATAAGATTA TCAAAGAAGT ATAATTGTCA TAATTGGTTA
47851 TATTTGTCTT TGATTTCCAC AAACAATAAA ATCACTTGCT CAGGTACTTG
47901 TAAAACTAA GGACTCAGTA ATACACTATA ATCTTAAGAG TATTTTAATC
47951 TCTTCACTGA AATCTCTCAA TATTTTCTTT TTAGCTAAAA AGAAATTATT

FIGURE 3P

48001 GAGTCCTCCC TGAGAATTG CTTTTGTAA AATTAGAAAT CATGTTTGCC
48051 ATTAGCGTTA GATTTTGATG GTGGGATAAT CTGGATATAT TCTACATTTT
48101 TTTCCCTCTG TTTTATGCCT CCTAACTCTG CTATTAATAA ATCATTCCCN
48151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
48201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
48251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
48301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
48351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
48401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
48451 NNNNNNCCAT TCCCATATTT ATTCCTCTA CATTGAGAAC TGTTTTGCTT
48501 TTAATAATTT TATAAGTTT ATTTGGAGTA GTGGTTATGT AATAGATAAA
48551 CATGAAAAAA ACCCAAAAGT CATAATTTAA CACACATTTT ATGTAACAGA
48601 AGATAATTTT TTAAATTTCC TCCGTAATAA ATTTACTGAA AACCATGGTG
48651 CATAAAATCA TGAGGATTAA AGGGTTTTTA AAGAGTCATG TAGTCATTCC
48701 CAGTCTGGAA TGTATCATTC AGGCAGATTA AATATTGATT CTTTTTGATG
48751 AAGATTATCA AGGAAAAAGA TTTCACAGAT GATCTTTGCA ACCTCTGAAA
48801 ATGTCTATTT AAAAGCAGTT GTATCTATCA GTAAAGGGAA ATAAAAACCG
48851 ACCATAAAAC AATATGTAGA ACATTTACTC AAGCTGTATT TATGCAAATT
48901 GAATTTATAT AAGTGTGGA TAAGAAAAGT ATTTGTCTTT AGACAATTCT
48951 GAACACAATT TTATAATATA CACGTAGCAT AGGGATACAA CGTAGATTCT
49001 AACTATGCAT TTGCTTATAT TAAAAAACT TAGANNNNNN NNNNNNNNNN
49051 NNNNTCTGTC AGTGTAAATG TTCACTTTTT AGGGAGGAAT GGTACATAC
49101 TGATTTAAGT CTTTGACTCT TTCAGTATAT CCCAAAAGTT ATGATTGAAC
49151 CTTCAGGTCC AATTCATGTA CCAAAGGGAA ATGGTACCAT TTAAAGGGT
49201 GACACAGGAA ATAATAATTG GATTCCTGAT GTTGAAGTA CTTGGTGGCC
49251 TCCGAAGACA CCATATATTC CTCCTATCAT TACCAACAGG CCTACTTCTA
49301 AGCCAACAAC AAGACCTACA CCAAAGCCAA CACCAATTCC TACTCCACCA
49351 CCACCACCAC CCTGCCAAC AGAGCTCAGA ACACCTCTAC CACCTACAAC
49401 CCCAGAAAGG CCAACCACCG GACTGACAAC TATAGCACC A GCTGCCAGTA
49451 CACCTCCAGG AGGGATTACA GTTGACAACA GGTACAGAC AGACCCCTCAG
49501 AAACCCAGAG GAGATGTGTT CAGTAAGTCT AATAAATGTT AGCACATTTT
49551 CAATAGGCTC TTTATAATGA CTTTCAACC ACAGGCCATG CCTTGAATAA
49601 GAATGAAACT CGTAAGAAGA ACTAGCTATG TAAAGTCGTA TGTCCCTATT
49651 GACAAATATT ATAAAGAGCT ACATAAAGAG TCAGTCTAAT TGGGCAAGTA
49701 AGAAAGAATG TATGTAGCAA TGGAAGGAAT ATTCACAAAG TCATATGGTA
49751 GATGACAGCT CTTAGCCAGT ATGGGAATTC TGACATAGTT GGATTTACTT
49801 GAAAACTCTC AGAGGTGGGA ACTTAACTGT ATTCATGCT ATGTTACTTT
49851 TAATCTAACC CTTCCNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
49901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
49951 NNNNNNNNNN NNNNNNNNGG TGGGAAAGAT GGTACTCAG AAGATACCAG
50001 TGGTAGTATG TTAGTGAAAG AGATCTGTGT AGGACACTAT GAGAGGAGAA
50051 TGAAGGGATA CCTCGTCAGT CTTTGGGTAG GGAATATTCA GGGAGGGCTT
50101 CCCGAGTGAA ACGGAGCAAA GTGGATGAGG AAAGGAAGAT ATTATAGTTA
50151 ACAGGGACAG CCCAACAAAT GCAAAGAAGT AGGAAGCAGT GTAATAGGCA
50201 GTAGACTGCC AGTAAATTGG AGGGTCATGT CAGGGACAGA TGAGAGAAGC
50251 TGGACAGCTA CCCCTTGGAT TGCTGTAAAC TGATAGTTAG ACCTACACTG
50301 GGGCTGGATT AGATCAAAGA CTCCCAGATT CTTAGCAGGT GACTAGGTAA
50351 AGGGTAGCAC CACCAATGGA GATAGAGGAT TCAGGAAGAG AAGCAAGTTC
50401 TGAAGGAGGA AAAGTTGTT AATTTATTAC CAGACCAAAA TTTGCATGTT
50451 ATAAC TTCA TCTGTTCCCT ATATTTATCT GCTTGTCTCG TGTCAC TTCTG
50501 CTTATTCTTA ATTACCTTTG CATTCACTCT TCCTGTCTCC TCCTTTATAC
50551 ATTACCAGTT TGTTTTGTTT CTTTCTCAT TCCTCTATGA TGAAGCTAAA
50601 GTTTGTGCCT TTTATATTGA ATACCTTTAT AACTTTTTAT CAGGTGAAAA
50651 AAATACTGAC ATTTATTATT CTTTAAAGGG GAAAANNNNN NNNNNNNNNN
50701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
50751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
50801 NNNNNNNNNN NNNNNNNNTA ATAAAGGGGA TATCAGAGCC GCATTTTGCT
50851 TGTATGTGAC TCTCAATTAA ATTTCTCCT CTGGGTAATT GTCCATTAA
50901 AGAAAGAGGG AAATCTCTA ATTCTGATAT AGTCATGTAT GTTTTGATA
50951 ATCAATATAA TTTTATAGGG AATTAGATGG TAGCTCTCAA AATATACATT

FIGURE 3Q

51001
51051
51101
51151
51201
51251
51301
51351
51401
51451
51501
51551
51601
51651
51701
51751
51801
51851
51901
51951
52001
52051
52101
52151
52201
52251
52301
52351
52401
52451
52501
52551
52601
52651
52701
52751
52801
52851
52901
52951
53001
53051
53101
53151
53201
53251
53301
53351
53401
53451
53501
53551
53601
53651
53701
53751
53801
53851
53901
53951

51001	TTATATGTGC	ATAACTATTT	TCTGAAATTT	ATGTTTTTGT	AATTTTCTCC
51051	TCACCCCTC	TCTCCATCTC	AGATACTGTC	TTTTTCTCCC	TATTACTCTC
51101	CTTTTTAAAT	TCTCATCATT	GTGATCATAC	AATGGGATTT	TTAATTTATG
51151	AGTGCTTAAG	TAATTATGGT	ATTTACATTA	TTTTGCTGCC	TTAGGATAAC
51201	AGGGAAATTT	GGCTATTTAA	TGTAAGATGA	TACGCTAAAT	ATTTTTTTCA
51251	TTATGATGAA	AGAATACATC	TTTCTGAGAA	TTTAAAAAAA	TCCTTTCTCT
51301	TTTTAAATG	TTTCTCTTTA	TTTCCAGTAT	TCCTCTCTAT	GCCAATACAT
51351	ATATGAATTA	AAAATGACAT	TGAAGTTCAC	CAATAAAAT	TAGTGCAATA
51401	AAATTGGGGA	AATACAGAGT	TCCAATGATG	TTGGGAGCA	TTCATTATAA
51451	GAGAGTGTGA	TCTTAAAGAC	ATGCTCTGGG	AAAGCATTGC	CTTGATCAAA
51501	TGCCAAGGCT	GTTGCATGCC	ACAGATAGCA	TTGCTGCCT	TAGAACTCC
51551	TGCCAAAAAT	TAAATTCCAC	TCTCATTTCA	TCTTCAGCGA	TAGCTGCTTA
51601	TTAAGTCTAG	CATGTGTGAG	AATGCTTTAG	ATGCTTTTTG	ACTTGCTGTC
51651	TTGTGGTTAT	AGCATATATT	CTAAATAAGG	CAAAGGCTCT	AAGTTTTTAA
51701	TTCAGGACAA	ATAACAGTG	TCTATTCACT	GTGAAAACAG	TAAGGAACCT
51751	TAGCACTATT	TCAGTTCAAC	ACCATTTCTT	TACTGAATAT	TTTCTTCACT
51801	TCTGGCATAA	CTTCTTAAAT	CATCCTCATT	CTCCGAGACA	GGTCTGAGAA
51851	TAGAATTTAG	CTCACCTCAT	TTCATAAATT	TTAGGTGAT	GTTAATTATG
51901	GCCACTATTT	GGCAATGCC	TTCAAATAGC	CCTCCACTTT	GTGTTTTTCT
51951	TTACATTGAA	AGATATTTTT	AAATTGANNN	NNNNNNNNNN	NNNNNNNNNN
52001	NNNNNNNNNN	NNNNNNNNNG	AAAGATTTCT	TTAATGGAA	GAGGTTGTGA
52051	AAGAGAGCTT	TTCTTATAA	TGTATGTGCG	TTGTTGCCCT	GAGAAAAGAC
52101	TGTAAAGATA	TTCTAAAAGA	AAATCAAGGA	AGAAAAATAT	TATAACAAGA
52151	ACACATCTTC	AGCCCAGACC	TCTCCCCCAA	ACTCTAGACC	TGGATGTCAG
52201	CTGACTGCTT	AACCTATCCA	GTTGGATGAA	AATAGACGTC	TCAAACCTAA
52251	CATGTATAGA	ATTGTTCCCT	TTCTTCCCTT	AAACATCCTC	TACTCTCAGC
52301	CTTTCTTATC	CCAGGTAATG	GGACCTCTAT	CACCTGTTGC	TAAGGTCTAA
52351	TAATGTGGAT	TCATCCTTAA	TTACTCTTCA	TACATACAGT	CAATATGTAA
52401	GGAAATCCTG	TTAGCTGTAC	CTTCAAATTA	TATTTAAGTG	TGACCTTTTC
52451	TCACCAACTC	CAGTGTACAC	ACCTTGGCCC	AAGCCATCTC	TCCCCTGGAA
52501	TACAGGCCCC	AAATCCTTTA	TCCAAATAAC	TTATGGAATA	TAGCATTTTT
52551	TAGATTTTAG	AAATCAGTA	AGGTACACAT	ACTATATAGT	ACACACTGAA
52601	GTAGTGAAAC	ATGCTGATTT	TCCTCTAGTG	CTTTTACTGT	GAACGTATCA
52651	ATGTTAAGGA	AAGGAAAATG	ATATTAAGTC	AGAGATGATG	TCTCACAGCA
52701	TATATTTACT	AGCTTGACAA	AATTTTTTAA	TGTTAGCAAG	ATTTAAGGCA
52751	AATTTTTTAC	TTATATTTTA	ATTGGATCCT	ATGATGATTA	TTAAAGAAAA
52801	AAGTAGTTAT	CTCTAGAAAG	TATAAATTGA	GCTCTTGGCA	AATGTGAGCA
52851	AAACCAGAAA	TCAGATTTTT	TTAAAGTTAC	GTGTACATTT	GTCTATAGAG
52901	TTATAATTAA	AAGTATTTGT	CTCATAGCCA	GTCTGTGATC	TGGGGCATAT
52951	TATTATATCT	TCCTATGCCT	TAAAACAATG	TTTCTGTAA	GTATAAACAA
53001	AGTACCATTG	ACACTGCAGT	TTTTGTCCAT	TTGCTACTTC	CAAAAGAAAG
53051	GGGTACAAAA	CAGAAAAGTT	ATTTAAACTT	TAAGCAGTTG	GTTAGATATG
53101	GTAGTTATAA	TGAGCTTTGG	TATTTTAATA	TGGGGCTTAA	AAATTTTCAT
53151	CTAAACCAAC	ATTAATTATC	TAAGTGTGAT	ATCCTTAGGA	GGTCTTTGTT
53201	AAGTCCTCCC	AATTATATCC	AGCATCTCCA	AAGGTGACTG	AAGTCCCTTC
53251	TAATAACTCA	TCNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
53951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN

FIGURE 3R

```

54001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
54051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
54101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
54151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
54201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
54251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
54301 NNNNNNNNNN NNNNNNNNNN NNTCATCCCA CCACTGGAGC ATTTTTCCTT
54351 ATATTTAACC TAAATTTATG AAACACAGT TTATGCATTT TTAATATTGT
54401 CAGATACTCA GGGTTAATGG ACACAGTTGT TCATTGTCTC CAGATCCCAT
54451 ACTCCATGTA CTCAAAGACA CTTTCAGGAG ATGACTCACT CTTCCCTGGG
54501 CCCNNNNNNN NNNNNNNNNN NNNNNNTTTC TTAGAAATGC TTTTTCCTCA
54551 CCCACTTAGC AATCTTGTG ACTGGTCCTT TGAACATTCT TTATTTTCCA
54601 AATCTGTCAC AAATGATCGA GTCAAAGATA TATAATAGTC AAATCTGACC
54651 CAATACTGAT CACGGGGGGG TTTCTTTTAA CATGTTTTGT AATCTAAATC
54701 ATTTCTCTTT GTTGCAAAGT AGAGATTTAT GGCAGCTAT TTAATCTCCA
54751 GATGTTTTTT CCTCAAANN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
54801 NNNNNNNNNN NNNTGGAATT TTATATTGGA TTTAATTTTT TTCATTCTAA
54851 GTAATCCTTC ATATGCTTAA CTAGCTAAAT TTTACTTGTT GGTTCGTGATA
54901 TTCCTTCCCC AAATAATCAA AGTTAATTCA GATTATGGTT TTTTTTCTTA
54951 TCCTGATACA GCTTAATATA GCCATTGAAA ATATTCTCTA GAAAATTCTA
55001 AGTCTAATAT ACTTCTAAAT TTTAAGCATG AACTGAAAA ATAGTAAAA
55051 GCATAGAAAA GGAATATGAA AATAGCATTG AAACATTTGG AATTAAGGGG
55101 TTTTAACCAC TACTAAAGAT AGAATATACT GAACCAGAAA ACCTAGTCTT
55151 GATTTCTACC TCTATTTATG AGGTACATGA CTTTGGACTA GTATGAGTNN
55201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
55251 NNNNNNNNNN NNNNNNNNNN NNNNNNTCCC AGCTATCCTT TCTCACTGAG
55301 TAGCTTCAAA GATCAGAGAT AAGGGATCTA GTGTTTGTGA ATTCTGAGTT
55351 ACTATGAAAT GTAAAAGATT AATGTGTTTT GTTTATCACT GTAAATAACC
55401 TTGATGGGGA GAACTTATGT GATTAAAGGT TGCAACTTTT TTTGTGTGAT
55451 GTGAAATATT AAACATATTT TTAATAAAN NNNNNNNNNN NNNNNNNNNN
55501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
55551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
55601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
55651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
55701 NNNNNNNNNN NNNNNNNNNN AAAACAATT TTAGCTTGAG GGACCCTAGG
55751 TTTTTTTAGG TTATATGACT AGAAATATAA TAAGATGTAT TTACAGGGCT
55801 AATCCGGTGG GCATGATTTG AGCAACACTA AAGTTATCTG ATGTAGCGTA
55851 AAATCTTGT TGTTCCTTAA ATGAGACAGA CTTGTGCTTC TCAAAGTGTG
55901 TTTTATAGAG CATTATTTCT CTGAAATATT ACACAGTTGT ATGTGTTAAA
55951 AAGTTCTTAG GTCAAATACA TTTAAGAACT CCTAGGTTTA GCTAAGTTAA
56001 ATAGGTTTCT TAACTGTAGG ACTTCTCAAT ACTCCAATGT GCTAACAATA
56051 TTGCAAATCT CTAAGAGAGG CATATAATGT ATCTTTTCCC AAGTTGTTT
56101 AACCATGGGA TTATTTGAAT CTTGGGATTA GAGTTTACTG GAATCAACTT
56151 TAGGTCAAGC TGGGCTAAAC TAGGAAGTAT GGAGACCATA CTCCACAAAA
56201 CCCAGTATCC CTCACATGTT AGCTCTCGTT GCCTGGTTTT TGAGACTAAT
56251 GGGTTGATCT TAGATTTAGG AATGAGGGTT GGGGAGGGCA GAAAGGTGCT
56301 GCACAAACTT AGCCAGAAAA GCAGCCATAA GTGGTGGTAT GCCTGTATGG
56351 ACACCATAGG TTTGCTGCAG GGGATTGTTA AAAAGCCACT TTTCAAAGAC
56401 AAAACTCTGG TGTACAGGG GTAGGAGCAA TCCCAGGGCA AAGAAACTCA
56451 GAAAAACTTT CTGCAGAACC CAAAGTAAAG GAAGTTAGGG GGACATGCTT
56501 TCTGTTCTAG CAATATATAA TGAAATCTTA TATAACTATT AAACATGTTT
56551 TATGAAAAAT ATTTAATAAT TTTGAAAAAT ACCCATGTTA CAAAAATGTT
56601 ACCTGAAAAG TAGAAAGCTG TACCTGCCAT TTGACTCCAA TTAGATATAA
56651 AATACATGTC TGCATATGTC TGGATGTGTG TAAAAGTCTG AAAGAAAAAA
56701 GACCAAAATA GTAACCATGA GTATATCCTG TTTATGATAT TTTGGGCGAT
56751 TTTCATTGTC TTTTGTGTGT CTTTATCCAG TAATTTTTTAA ACTGTTCAAC
56801 TGTACTATCG TTGATAACAG TAATTCATTA TATCATTAAT TCCTTCAGAG
56851 AAGGGGCAGA GTAACGAGAC CCCACTCTTC TCAGCAGGAA AGTGAGGATT
56901 CTGATTCTTA ATCATGTTTA TTTTATTTTG GACATACTGA GTTTTAAAAA
56951 TTCATTGAAG GTTTTATTAA CTTCTATTTT TAACCTTTCA GATTTTCCTT

```

FIGURE 3S

57001 TACAATTGAA TGTGATAGAT ATGTCCAATT TTAGCTAGGA AATTTGTATT
57051 ACGAGAGGGA ATATAACATC ACTTGACCA GGAACAGAT TTTTTTCCC
57101 CAGTGTGTGTC AGAAGCCCAT TTAACTAAA ATAACATAGG ATGTTTATTA
57151 ACTAAAATAA CATAGGATGA TTATCTTAAG ATGTAGATTG CTAATCCTTA
57201 TGACACATTT ACTAAATCAG ACTTACTAGG GAAAGTATCT CTCAGTCTTT
57251 ATCNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57451 NNNNNNTCTG AGTAGGCACA CTTAGTCCCC AACACTCAAC CTTTTATTCA
57501 ACCAGCGGTG AGCTGGATGT GAACATGACA GACCCAGTAG GGTTCCAATG
57551 CCTGACAACC TGCACCTGTC AGGAAGAGCC CCCTCCTTTC TGCTCCCCT
57601 GCAACACATG GTTANNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
57851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN ATAGCAGGC TCTTCTAAA
57901 GCATAATGGC AGCTCCAACC CAGCAGAACC ATATTCTGGA TGGAGTCAGA
57951 ACTAGCCAAA TGTAGGCTAT TAAAAATATA GCAATTGTGC CTAGGCATGT
58001 GTCCACCTGA CTCATGTAAA ATAAAAAGAA AAGGCGTTAA TAGAAGCCAG
58051 ATTATAGAAC ATAATGTTAT TAATATGTAT CAGGTGCTCT TATGAGTGGC
58101 CATATCAGAG AACCAGCCTC ATTGTGTGCTG TTATTACGCA ACATGAGACT
58151 GTGCTGCAAT TCCAAGTACC AAGCAGTAAG AAGAAAATTG TTTTCTTTTA
58201 CTGATTGCTG CATGTTGTCC TGAGGTTTTT CCCCTCATCT CTCTTCTACA
58251 CATTTTACAG AGTCTGTGGC TAGTTCAGTC ACATTGTTCT TAGTCATGGA
58301 AATATTATGG TCTTCTGAT TCTGTGTAGT GATAAGTAAA AAGATTGTTT
58351 CTGCTGAGGG GTGAAAAGGT CTTCAAAGTA GTTTGCTTTC TTGTAAACAG
58401 TAAGACCTGC AGAGACCCTG AGAAGATGCC TGATATCTCC TTGAAAATTA
58451 AATTTCTGCT AGTGTTTTGA AGGAGCGAAT TGTCACCTCT CACAGGTTAG
58501 GATCTGCTGC TGTGAATTCT GAAAGTTTTC AAGATTTTTC ATTTATATTT
58551 TAATTAGATG CATGTGTAGC TAGTAAAGAA AATATCATTC TCTCCAAAAC
58601 GTATAAATGG GCCCTTGGTG ATTGTGACCA AAGTCAGAAG TCAGATTTT
58651 TTTTGAAGTT ACATGCACAT TTGTCAATAG AGTTATAATT TACAAGTATT
58701 GTCCTCANNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
58751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
58801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
58851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
58901 TATTGATGTT GAATTCATTT ATATAGTAAT GATAACATTT CCTACTAAT
58951 TCATAAAAAG ACAGCCTATG CTGTTTTCTT GTTCTGAGTT TATATGTTTC
59001 TCATGCTTTT TATTATGGTT CATTACAATT TTAATGTTAT TTTTAACTAA
59051 CTAGATCCTT TTGAAACAAA TTGGTTTGCA AGTGTGAGCT GTTAGGTGCA
59101 CAGAGAAAAA TGAAAATAGA AACTTGCGAT TTTATTCTAG GCTTGTACC
59151 AAATATTTAG AATACTGTGT TTTATTTAGG TGTTTATAGT CTCATTAGAC
59201 AGTTGTGATT TTAAAATAGA GACCACATCA TCTCAACTTC TTTACTGTGA
59251 AAATAATGAC AATAGTCTTT TCAGAGATGA ATCTGTCTAG ATGGGAAATT
59301 TACATGATTG ATCTGATGAG NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
59351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNAA ATAATCTGAC AAGTAGTTTC
59401 CCCAGAAAAT CTGATTTAGT AAATGTACCA AAAGGATTTA GAAATCTACA
59451 TCATAAATAA ACATTCTATG TTATTTTAGT TCAGACCCTA TTTTAATTCA
59501 GACTTCCTAT GGGATAAAAA CTTCATTCT TTCTTTAAAT AGATTCTTTT
59551 GGCTTGAGTG CATTTACACC TGTCCTCAAC AGCTGGTGGG CTTCTGCTCA
59601 CCCTAGACGG TGTTCATGCT GCACTCAGTC CAAGCAGCCC TTATCAGAGA
59651 GTCTTCTTAC CACTTGCATT CTGGTGCGAA GGACTCATTC CTGGCAGAGC
59701 CTAACCTTCA TGGGAACCAT TGGCTTAGAA GAGAGGAGAA GNNNNNNNNN
59751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
59801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
59851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
59901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
59951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNTATCC

FIGURE 3T

60001 TGTCTCCGGG CAGTGTCTCT GAGTCTTTCT AAAACTTTAT TAAAAGTTCC
60051 TACAATACAT AAAAGAGAAA TAGGTATGTC AAAAAAATTG GGCTCATGTG
60101 TGATTAGCAA TGTCTATGCT CCACCTTTCA CCAGAGATTT AACTTTTTTG
60151 CCAATTTTGC TCTTTACTGT CAGCCCAAAG GTGTCTGTGG AACTTTGTAG
60201 ATTCTTCCTA TGGGATGAAA GCTTTGACAA ACAAGGTCTT ATTTCTGTGA
60251 CTCCAACAAC CCTTTTCTTT TGGCTTATAA TGAATAGATG GCTAAAACTT
60301 TTCTCTAACA TGTTTAAAGA AATTTCAGAA ACTAATTTCC AGCCCTTTTA
60351 TTTGCTTTTC TCTCTGATCA CTTAAAATTG TGATGCATAT GTCCTGCATC
60401 TGTNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNCATC
60451 TCATAAAACT CTTGTCATTC ATCCCTACGC TCCAATCTGT GTCTCTAAAA
60501 GCCTCTAGGT TTTGCCACCA GAAACAGCCT TCAGAATATT GGAATAATTA
60551 CATATGTACC ATTCTCTTCA ATGAACTAAA CTCCTAAGCA TAGAAGTAAT
60601 TTAAGACAGT GTTATTTTAA ATCATTTAGT CCAACAACAA ATGCCAGAAC
60651 TGGCCTGGGT AATGCCATAG GAACCACTT GTTTCCTATG AAACAGGAAG
60701 AGAATTTGCA AACCTACCTT TACTAAATAA TGCTTACATT TTGCTATAGT
60751 TACTCTTGGC AAGAAGTTGA GCAGTGGGAG GTGTATATGA TAATATTAC
60801 ATTTACTTCT CTGTGCTTAC TGTTAATGTT TTCTGGGTAA AAATATGCAA
60851 TTGACTATT TGGAACTTC ATTTGTGAAC ACAGGTTATA TATGATCACA
60901 TCCCTGGGGA AATAACATAT TTAGTAGAAA GTGCCTGGCC AGCATTTCTT
60951 ACAAGAACTT TCCTGCTCCT TTTTGTGGCT GATTTTACCC CTGACTCCAG
61001 GGCCCACTAG CCATTAGAAA GTACTGTGCT CCTCAGCCCA GTCCAACCAT
61051 GCCTTTATCC AGCTGTCACT TGCTTCAGTA CCTGTACATT CCCTTTCTTT
61101 TAGTTTATGC AGAAGCTGT AAGAAGCAAC AGGCAAGACA TATCTTTTGG
61151 TGGGTAAAC ATGGACCAAT GGTATAAAGA TTCTGGGGAC ATTTTCTAAA
61201 AATATATGAA CATTATGGT TGATGTTAAT TTTAGTCATA ATTCCCTAT
61251 GACACCCACT CTCACCTTCT TCCTGTTTCT TCTCCAAGCT ACTACCAGTA
61301 ATGGTCATTT CAGCATGCAA ACCTGATCTT GGCAATCCTC TGTTAATCCT
61351 TCAGAGGCCG CCAGTTTCCA CCAGGGATAA CGTCCAAAT ATTTACAAGG
61401 CTTTAAAGC TCCTGCACAC TCTGCTTCTA TCTCCAATC CATTGTATGC
61451 CCTTCCTCCT GCACCTTCTT GCTGCCTCCC TCCCCCTGTA AGGTCTACAG
61501 CAATTCTGAT GTTTCAGTCC CCTCTCTTTT CATTCTCCAG AGCCATAGCT
61551 CATAGTACTT CTTCTTATGT CATAGAGAAA GCATCACTT TCAGAGATT
61601 TTATGACGTG TATCCTCCCC AGCCTAGCTT AGCCAGTCCC TCACTCCCCA
61651 CTCCCGTACT CACTAGCATC CTCTGCTTCT TAAAGCACA CATTACCTGG
61701 GGAAATTGTT AAAAATCNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
61751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
61801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNAATGAGT GGATGGAGTA
61851 GAAGGAGGCA GGATGCATGG ACAGATGGTG GAGGAATTGT ATTGCTAGCA
61901 CTTCTCTGAT TCTTTTATAG TTTCTAAAT TATACCATT TAATTCTAAG
61951 AATTAATGGT TCAACTGCAG AGAATAGCAG GTATGCATTT GTTTTTATGA
62001 AATATATATA TTGTGTATTT AAGTTGAATA TTAGCCCAAA GTATCAAGCA
62051 GGGGAAAGAA CAATTAGAAA AATCATGTAC CAATGTCTAT TCCTGGGNNN
62101 NNNNNNNNNN NNNNNNNNTT AATTAACATT TTCTGCATTG ATCAAATAG
62151 CTCCCTTGAA ATCAAAGAAA AGTGTGTTGAA TTTCACAAAC ATATTTCAAT
62201 TTCATAGCCT TGCGTTCAAC AAGTATCAAG CCTATTCTAA GTTCTCTTTG
62251 AGTAACCAAA ATACAAATAA CAGACATATA ATTGCTATTT ATATGTGATC
62301 TAACAGAACC CTCTTTTATT TAGATGTCTT AGGTAAGTTT TTTATTCATA
62351 TTTATATTTT TTCTTTTAC AATCCTAAAA ACATAGAATT AAAAAAAG
62401 TGAACCAATA CAAATGAAAA AAAAAATTTT ATTTCTAATA CTTAGATTAC
62451 CAATATTTAG CATATTGGCA TATTGGTGT TTTATCCCTC TTTAGTCCTT
62501 TTTTTTTGGT TTGTGTTTCT TGTACATGAT TGTGATAATA ACATAATACT
62551 GCTTGAATGA TTTTGTCTGC TACTTATTTT ACTTAATATT TTCACATAAA
62601 CTCTAAAGTG TATCGGGGGT GGGTAGTTTT TTTTCTCTTG AAACCTCAGT
62651 GGGATGCTCT TAGTAATCCC ATACTGGTAT GTGTGAGGAA GGAAATTAG
62701 TTAAATAATT TGGTATGGTT ATAGAGTAGA GCACACAAA TTGTAAGAAC
62751 CAATAGCTTC TGAGCACTTC TGGTCTCTAA ATTCCTTGGA ATGTTTCCCA
62801 GTGGATTGTA ATGAAGGTAT ACATGATCAT CTGCTGCTAA ATTAAATGGT
62851 TCTTAGAAAC CAAGAACCNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
62901 NNNCATCTGA CATTTTACAA GTTGTACTT GACTAATTCT TTGTGTTGCC
62951 ATCTTCTAAC TNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN

FIGURE 3U

63001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
63051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
63101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
63151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
63201 NNNNNNNACA GGCTCCCGGT GCCAGGAAAC CTTTACATCA ACACTCGATT
63251 TGCCATTGTA TAGTCCTTCA TCTGGGAGGA AAAAAAAGA CGGAGGGGAG
63301 CTTGAAAAAC TGTCAATATG TCCCTGGAAT ATGGTACTTT TAAGAGTTGA
63351 GCCTATTCCA TTTTGGAGAT GATTATATA AGTTACAACA AAAGAAGGGG
63401 ACAAAAACAT GATTGTCTA TGGAGTTTTT ATAACTTTCT GTCACAAGAA
63451 AGCACGCTTG TCTACAATTT TGTAAATATTT CTAGTAAATA AAAGAGGCAC
63501 TCCCCGTCTC AGAGCACCAA ATAAGGAAAG TGTAATTGGA TGTCAATTGCT
63551 GTCAGTCAGC TGGGTCTATA AAGAGAGAGT GGGGTGCGCT CATCCCCTGG
63601 GTATCCACAG TCAGCTGTGT CCCTAGAGCT TCTTTTCTTT CATTGCTGCC
63651 CAGCTGGGTA TATTGCAAGT ATGGATTATA AGAGGGGAAG GGACTTCACT
63701 GTTTTAAACGT TTGAAACAAA AAGGAAAAAA CTCAGAAGTA GTAAGCTAAA
63751 AACAACTTGT GCAAACGTTT TGGGATTATT ACTTAATTTT AAGAATTTTT
63801 GCTAAAAACA ATAGGAACAT CGTTGAAAAA AGAACCCTTT TGAGTGATGA
63851 CTGTTTCATA GTGTTCTACA TCTCCTACTC CCTGCCTTAT AAAAATAAAG
63901 CTAATTAATT GAAGTCTGG CAAAGAGAAG GGAGTATTCC TTGGCCTTTG
63951 ACCANNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
64001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
64051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
64101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
64151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
64201 NNNNNNNNNN NNNNNNNNNN NNAATGATTT AAAGAAATCA ATTATTTTTT
64251 TAAAGTAGAT TCTTATGCTT TGTCCCACTT GTGTCCTTTG AGGAAAAGTG
64301 TGAAATACTA CCTGGGGTTC TAACAGGATC TTTGNNNNNN NNNNNNNNNN
64351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
64401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
64451 CACATGTGTG TAAATAGCTG CAATTCTTTA TGAAAGTGTG TAATAGGGGC
64501 TTATAAGCAT GAATCCTTGA GGTACAAGAG GGAAGAATTT GATTCTGCCT
64551 GGGGCAGAAAT CACTCCACAG AGGCCGTGAT ATTTATGCTG GGCCTTGAAG
64601 GAAGAAGAGA TTTTGGCAG GTTAATTAGC AGGAAAGAGA GCATTGGGTC
64651 TGAGACGTGA AAATGAATGG CATGTTTGGA CAGCATTGTG TATTTTGGTG
64701 GGGTAGATTG TAACTCTGG CCAAGAAGT TTGAGTTTGA TCTTAAAGGA
64751 GCCAGGGAAG CTGTTTAAAGC AAGGGAGGCC ATGGTAGGAA ATGTATATTA
64801 GAGTGCAACA ATGTGCGGAG TTCCTGGGG GTGGGGAGTA CATGCTGTGA
64851 TTAGAAACCA AACCATTCGC TTCCATCATA GGTATTTTCA GTTAAACAA
64901 CCTTGATTTA AAGATCATAA TAATGTATCT TTCGTAGAGA TGAACCTTAG
64951 AATTTTCATAG AGATGAATAG TTGGCTAGAA ATTGGTTGCT CTATTAAAT
65001 GTATTTTCCA GATAAAGACT TTTTAAGTTA TTTTGTGTTA ATTTTAGAAA
65051 CTTTTGAAAA TCATTTTAA AAAATGTACC CAGATCAAGG TTGCAGTATA
65101 AATCACTGCA GAACAAAATT CGTGAATCTT CATGTGATAT ATAATTAAAT
65151 ATTTTGGATT CATCCAGATG AATAAAATGA CTAGGCCTTT TCCTCTGCC
65201 CTTGTTAAAG GTAATGCCTC TTTATCAATA CTTACACAC ACACACAC
65251 AAAAAGTCAA AACCTTTCTC TTGCAGTTCC ACGGCAACCT TCAAATGACT
65301 TGTTTGAAAT ATTTGAAATA GAAAGAGGAG TCAGTGCAGA CGATGAAGCA
65351 AAGGATGATC CAGGTGACAC TTACACTCTT AGTGCCATCA TAATGACATT
65401 TTATTGTTTC TCTCCATGAA AATAACTTTT AAATGTAGAT CGTTTGGACC
65451 ATTTGGGAAA ATTACACCTG TTTTCTTTAA CAACTCAGAT TTTCTTTGTA
65501 GTTAAATCAG TTGAGCNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
65551 NNNNNNNNNN TAGTACTTAA ACTGTAAAAA GGAAATCATA TTTTATAAGT
65601 GATGATGTAA CATTTGTATA ATTTGGCAAC ACTTGTGTGT TCTATTTTAA
65651 CAGATGATGC ACTATTGTGC TGAGAAAAGA TTAATGTAAT TTCCCTTTGG
65701 TCTTATGGAA AGAAGCTCTG GTTATATGGA TAGGAGAGAA CACTTGTGT
65751 CTAGATACCC CTAGTTCATA GATTCATTTG TTGTTTTTTG TTGATTTTCT
65801 GATCCACTCA TCTTTCTGAT TATTTCTTTG TAGGTGTTCT GGTACACAGT
65851 TGAATTTTG ACCATGGACT TTGTGGATGG ATCAGGGAGA AAGACAATGA
65901 CTTGCACTGG GAACCAATCA GGGACCCAGC AGGTAAAACC ATTTCAATTA
65951 ACTTTTTCTG GTACACATTT CAATGTGATA CTATCTGAAG ACTCCACTGC

FIGURE 3V

66001 TGCTAATCAA GTCTACTGTA ACACATCTCT GTGTTTACTT GATAAAGATG
66051 GTCAAAAAGG ACCATGCCTA GCATTTATAT TTCATTTCTT TGAAAGAGTT
66101 AAATTTGTTA GCTATCAACT AGCACTATAT TATGGGAGAC AAGTAGTTAA
66151 TTAAAAGGTC AACTAATTC TCTCAAAAAG CTTGATAAAT AAAATATTTA
66201 GATTAATTCCT TGCAAATAAT TCTTAAGTTA TTTGTAGCAC CCATCCCAG
66251 GAATAAAGA AGTAATATAG TGGTATATCA GCCATGGTAA TAGGCATTCC
66301 CCAGTTGNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNAGAT
66351 TTGAGTGGTA GTAGAAGAAC ACAGTGAAA TTTCTATAAA GTTGTGATTA
66401 TCCTCTGAAA AAATTTCACT CTAAGTGGG TGCTGAAGGA AATGGTTACG
66451 AATCACCACC ACTCATTTTA TAACAAAAGG TAGATTTAGA CATAGTTCCT
66501 TGGAACTCAA AGGAGTTTAA AATGCGACAA TTCCTATTAA TCTAAAGTTT
66551 TAGTAATTCAT AAGACATCTT TTCTTCTAGA TTCAAGATAG CTTTCTCCC
66601 CTCAGTAGTT AAATCNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
66901 NNNNNNNNNN NNNNNNGAAG TTAAATCTTT TATGGATAAG GCCTACTTAA
66951 TGGAAAGGGTG AGGAGTTTGG TTTTCTTTT TTTCTTTT ATACTCCTGA
67001 TTCTAGATTT TAAGAAGATA CAGAATGTGA AGGAAAAGCT TGTTATTTCA
67051 TTTATTTTAA AAGCCTAGTT CTTGAATACC ATACTGGATA AGAGCACCAA
67101 ACACAGAAAA GTTAACTTGA GAGAGCTCAT GATATCTTAA ACCAGCTCNN
67151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
67201 NNNNNNTCATG CATTCTTAGC AGGCTCCCTG AAGATCTCTA TGGTGGTCCT
67251 CAACGCACAT TTTGAGTAAC AAGGTATTAA AAAATTAGTN NNNNNNNNNN
67301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
67351 NNNNNNNNNN NNNNATTTAC TTAAAAATTA AAATGTGAAT TAAATTCCTT
67401 AAAATAGGAA TATTCAATAA ATTCCAAATT GGTTCATACA TTTCTTTTCA
67451 TTGTAAGAAGA CAATATCATT TAGTGCCTCA CACACAGTAC TCTCAGTCAT
67501 TCTTGCATTA CTGCTTACT CATTTAAAGT TAACTCTGAG CAGGACAGCA
67551 TTATCATTTAA TGTACAGAT TCTGTGAATA AGAGAGAAAC CAACATATCT
67601 ACNNNNNNNN NNNNNNNNNN NNNNNNTACAA TATAAAATAC CTATGACAGT
67651 CCTAGCTTAT TTAAATCTC ATCAATTTTA TTTAATTGTA GCAATTAAAA
67701 TAAATTTTATT TCATTGCTAC AGAGTTGTAG CAATTTACAC TCCTACCAAC
67751 AGTATTTGAA AACCATCCCC AACAGATATT ATAAAAGTAT TTTTAAAT
67801 CTCTACCAAT CTTATAGGTG AAAAGGAATC TCTTTTAAAT TTACTTTGCA
67851 TTTCTTCTAT TGTGAAAAGT TGATCATGTT TTATCATTTA CTCGAGGAGC
67901 AAANNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
67951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
68001 NNNNCTCTGT ACTACTCATG CAGGCTCCTA GACAGTCCAA TACTCAGAAC
68051 ATCTGCAGCA TTATTGTGTT TTTGACTGTA CCACAGGCTG GCCTCAGTAG
68101 GTGTGCATGT GGTAGTGTGC ATGATCTGAT TAGCAGATCA CAGACCTGTA
68151 GATTAGACTG ACTGACCACT GACCTGGAGT AGTTTGCTAC AGAAGGAGGG
68201 ATCTTTGGCA GTCATGTATG ACAGAATGGT GAATAATATT TATGGTCATA
68251 ATGAATGATT CCTTCAAGTA AGTTACACTG GGGAGAGTAA TTCATGCTAA
68301 ACTCAGGAGA ATTCAGTCA CTTCTATGTA AACAATCTAT GTAAAATACT
68351 TGAGTTTTAG AATTTAAATA CTGTATTTTA GAACCAATTT AGTCCTTTTC
68401 AACATTTTTT TATTATTCAT TATATTTGTA AATGTTTAAA TTTGTTGTGA
68451 CCCAAATCTC CTTTCTCTC TGATGGAAAA ATGAGAAGAG GGTTAGAAAT
68501 TAGAATGGTA GACAAAGAAA TTTGAAATCG TTATTGGACC GTGTTAGTAC
68551 AGAGCAAATT GAAAGGAAGA GATGCTGCT TGAGTGATTG GATACTCGT
68601 GATGATTGGC TACGGTCCCT GGAGGATGCT TTTCTAGAAT TGGTCCAAGT
68651 TTTTACCAAG AACTTAGATA AATAGGACAT TGAACCTAAT ATTGAAGATA
68701 AACTGACCAA ACTACGTTG ACTCAAATCT GGAGAGGAAA GTGACTATGT
68751 GTGCTGATAC AACCAGCATT CGAAAAATTA GATTAATGGG TCACAATTAA
68801 GAAGGTGAAA TGTAATGGGG TTGAGTCGAA GTTATGGCTC ATAGTCGAGA
68851 AACCAGTCAC ACATCTGAGT GAGGGGAAG ACCTGACATA TCACTAAGGG
68901 TAGAAGACTT TAGAGTTTTA GTGTGTAATG AAATCAATCT AAGCGAACCT
68951 TGTGATGTTG CTGCCAGAAA TATGGACCAT CTCTATAGAA ATATGCTATG

FIGURE 3W

```

69001 TTCAAGAAAT GAAAATCCAC TTTGGTTAAG TGGTGAATCC ACTTTTGAGA
69051 ATTCTTTTTT TCTAATATGA GAAGAATAGA GTTGGTCACT TAGAAGAGGA
69101 TGACCTGCAT GGAAGGTAT CTCCAAGCTT GTGAAAAACA GGCAAAGGAA
69151 CCATAACGTT TAGCCTGGAG AAGAAATGAT GCATCAGGCT TCCTTTTCAA
69201 ATAATTGACT AGGGTTAGAG TTGTTCTCTG TGACCGAAAG GGGATGAGAA
69251 TAGGATTGAT GGGTAAAAAA TACAGGGAGA TTCATTTTGG CTCACTGTCA
69301 TCTAAAAAGA GAACAGGCTG ACATGGAAAG AGTGAGTTAT CTTGGTGGAA
69351 GCATTTGACT GGGTGACATA TGGCAAGATA CTGTTAAAAG GTTTTGAGCA
69401 TCAGATGAAA GTTGGATTAG AAAACCTCTA AACCAGCTCT GAGAGCCTGA
69451 TTCAGAAATGG AATGTGGAAA TAGTCCCAAG AAATTCAGTG AGCAGGATTC
69501 TAGCTTGGAA AGAAGGATGG GAAGACCAA AGATGGCAGA GAGTTGGAAG
69551 GCAGGATATG TGAGGTAGGA ATTGCAGAA ATATTTTATG AATATGTTGA
69601 GAGAATTATA ACCCAGCACT GATGGGAAAA TACTAATAAA GCTTACCTTT
69651 TCATGGTTTT TTATTCTAGG AAATTCAGGT GGATGGANNN NNNNNNNNNN
69701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNTC
69751 ACATGTGAAG GCCAGAGAAG AATATAAAAA ATCNNNNNNN NNNNNNNNNN
69801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
69851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
69901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
69951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
70001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
70051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN AATAATTTCT GTGTATTATT
70101 CTAATTACCA AGCTAATTAT TTACCTATTT CTATCTAGAT TATACCACTT
70151 ATAAAAGAAT ATGTGTTTGG ATATCCAATT ATGGATGATT TTAATGATGT
70201 GCAGTTCTAA TCCAAGATCC AGATTGTGTA ATATTCTTAT ATGCTAATAT
70251 TTCTAAGAAA ACTCTCAAAA CTCAAACCT TAAGAAATAG CTGCAAAATA
70301 AGTGCATTTG CTAGCTGCCT TTCATGGTGC TATTAGGTTT TATCACATTC
70351 AGTCACAGTG AGATGTGAAA TTAACAATGC CTCTAAGAAA TGGAAAATTG
70401 TGCTGCTGAC AGACATGGTG ATGCGCTGCA GGTATGTAGC TGTGGCCCTG
70451 GGGACACACC CTCTCTCTGT GTGTGTCAA GCTAGAACTA AGGCCCTTTT
70501 CCTGGAACCT TGATCTCTGG AGAACAAGGA TTAAGGGCAA CTGACTCAGC
70551 ACTCTGTCTG CCACTGATAG CTACAAGGAA TTTGCTTACA TATGTATCAA
70601 GTATGCCAGC ATGTACATAC AAGTATGTCT ATTTCAGCAT TAAAACATTA
70651 TTTTAAATA CACATAAATG TAAAATAAAT CAACAATATA AGTGAAAGCT
70701 TTCTACCACC CCTGACTCCC CATTACCAC CCAGCAAGAT GCAGCCAGTT
70751 TCTTGATTAT GTTCTCAGAG AGATATTTTT GGATACAATG GCATTCTCTT
70801 CTGTGCCTTG CCTTTGTCAC TTATCGATAC TACCTTAAGA GTTAACAGGN
70851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
70901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNTGGG ATAATAAGTA
70951 CTATCCACTA NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
71001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNGC TGAAGCATGT CCCCTGTCAG
71051 TACATACAGA ACACCNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
71101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
71151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
71201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
71251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
71301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
71351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
71401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
71451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNTCAT
71501 TTTTCTCTG ATGTTCTTTA TTTTCTTTA GTATTTTAT TGATTTATGG
71551 AAGTTTTCTA TACAAAAAGA AAAATTATCC TTTTTAATT TGTGCAAAAT
71601 ATTTAAAAAC ATTTTGTGTT GTCTATTTTA TGGTATTTT ATCTTAACAT
71651 GCAAATTTGG AAACTTAAAT GTTCTTCAGT AAGAAATTGG CTAAATGGTT
71701 ATTGGCATAC TTGCAGCAAT GCAAAGTGTA TTTGCTACCA TAGAAAAATG
71751 GTTATACAAG AAGCAACTTA CAACACTGCA AATATAGTAT TCCTTTTTTTT
71801 AAAAAAAGTT CATATTCATG TGCATTTATG TGAATGTATT ACATAAGAAA
71851 AGTTCAGAAG TCAAAACCTT CACAAATACC TCTAGATTT GTGAATTGAA
71901 AATTTAAAGG CACCAACCAC TACAGAATTA AAAATAAATA AACTAACATT
71951 TTATAACCCA ACTTTCTGGT CAAGTTTTCG AGCATCTATT ATTTACCCTC

```

FIGURE 3X

72001 CCTCCCCAAA ATTCTTGAGA TATTGTGATA CTGTAAACTG AATGAATGTT
72051 TAATATAATT GTACATTTAT CACCAGAATT ATGAAATCAG AAGTGAAAAT
72101 TTTAATAGGT AACTTGCTTG AACTGAATC AATTTTCCC TTAGTTAATG
72151 TTCCAACCTG TATCATTTCC TTTATTACAA AGGGAACGAA TTCTTCATGG
72201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
72251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
72301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
72351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
72401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNTAC
72451 CTGTGATGAT TCAAAAGCTG ACTTATAGTG AGACATGTAA ATATTTTGGC
72501 CCAAAATATT GTTTGGATTA GTTTGTGGAG GAAGGGGACA AGGGAAGTGG
72551 GAGAAAGATT CTTAAGACAT TATGTATTTT TAGTGATTCA AGTACGTGAA
72601 TCAAGGATCT ATTTCCCTGC AGTTTTTTGC TGTATCCTTG GAACATCATT
72651 CGTTGATTCA AACATCCTTA AACACCTGCT GCCTGCTAGC AGGGTAACAC
72701 AGAGGTTTGG GGGAAACTTT GGTTCAGACT TTTAAGTTGT TCACAACCTG
72751 CTGGGGAGCC CTTAGCACAG AGGCTGGCAC ACAGTAAGTG CACAGAAATC
72801 ACCTGTGACC TATAGAGCTC CGACGGGACT TGTTAGAGAA GACACTGGAA
72851 TAAAAAGTTA ATACCAAAAT TTATCATGTC CTGTATAAAT TAAGCATAAG
72901 AAAGAGATGA TAGTTTTATAT CTGTTAACTG AAAACTAAAT GGGATTACTA
72951 CCAAAATAGA TACACATGTT CAACATTCCA ACCATTTTTC TATGGAAAAA
73001 CCAAGTGTGA AACATACAGT GTGTTTTCAG TTGACTGAGA TAAAATAGCC
73051 TGAGGGAATG TCATTTGTAC ATGGATCAAG TTTCAAAATA CTACTTGCAA
73101 CTTTCTCTCT CTTGAGAAGC AGCACCCCCC ACTTCACTGC TGCTTAACCA
73151 TTTTTTTTCT AATGCAGTTT CCTTAACAAC AGAGGGAAGA AAACCTCTGA
73201 GCCCTAGGGA TCTGAGAACA TCATCAAAGT GGTCTATCTA AAGTGAAAGT
73251 TTTTATTAAT ATTTGTGTTC ATGTATATAT AATAGGACAT ATTTATAGAA
73301 TAGATATGAT ATAGATTTGT TTTACTATAA AAATGTTTTA AATGGCTTAC
73351 CTCCAAATAA AATTGAAGCT CCAGGAAGAC GGGGAGGGTT TCCTGAATAC
73401 CTCTGTATCT TCCTGGTGTT AGATTTACAC CAGTGTGAGA AGCTCTGCCA
73451 ACTCTGATTT TTTTTTAAA GAACCTTTAC GTTTAGAGAT TTTTTTTTTT
73501 TAAGTTGGGG GAGGAGAGAA TAAGCAGATA TAGGGCCCTT TAAATCACCA
73551 TCTGCTTTAC TTTTCAGAGG TTAATTATCC ACCAGAAAAT TCTCCCTAGA
73601 AATTTAGGGT CAGTTTTACC AAACCACCAA GCCAGATCGC CAGGGGTGTT
73651 GGACAACGTT ATCATAAACT GGTGAGGGTC ACAAGGTGGT GAATGTGTAT
73701 ATGTTGTTCG ATGCCTATAT GTCTGATCAT ATTCCATTGG AAATATTCAT
73751 AGTGTTTGGT TACTAATTAT TGTGTGTCTC TGGCATTACT TGTACTTTCC
73801 ATGAGCAAAG TAAGTGAATG TTTTCTGTAA ACACATACCA TAGAACTTAA
73851 GACAGGCATA CAAATAAGTC ATATTTTCT TCTTCAAAA TCTGGGAATT
73901 CTTTACAAAG AAAAGAAAAA AAAAAAAAAA CTAAAACTCC TGTCCATGCT
73951 GCTTGTCTA GAAAGGGATA TTGGCACATA CACACAGCCC TGCTTTTCCC
74001 ATTCCAGGTG GACAATATCT GACAGTGTCG GCAGCCAAAG CCCCAGGGGG
74051 AAAAGCTGCA CGCTTGGTGC TACCTCTCGG CCGCCTCATG CATTACAGGG
74101 ACCTGTGCCT GTCATTCAAG CACAAGGTGA CGGGGCTGCA CTCTGGCACA
74151 CTCCAGGTGT TTGTGAGAAA ACACGGTGCC CACGGAGCAG CCCTGTGGGG
74201 AAGAAATGGT GGCCATGGCT GGAGGCAAAC ACAGATCACC TTGCGAGGGG
74251 CTGACATCAA GAGCGTAAGT AGATCCACAA AGGAGGCAGG ACCTGGGACG
74301 TTTTCCTTTC ATAGGAGAAC TCTGGGATCT GAATTTGAAG AAGCCTTGCT
74351 GTGTGAATTC AGGCTCAGAT AAAGGTTTGG GTTTTTTCT GGGCCTGATG
74401 ACTCCATTCA GTGTCTCTTT CTGAAACCAC TTTCTGCTTC TTCATCCTTC
74451 TTCATTGCAA CCTGTGATCA CCACCTCTCC CTAACCCTGG TAAAGATCCA
74501 TTTTTTTTTA ATGGAAAAA AAATCTAGGA ATAGAAAAGT TTTTTTGTT
74551 TTTTTTAAAG TTGATCAACA GAACTTCCA AGGGTGGGGT GGTGAGCTTA
74601 ATTGTTCCCTG CTTTCCCTTT CCTTTTAATT CCTTAGATTT CAAACTTTTA
74651 CTACCATAAA TTACCCAGGA GATGGTCTTT TTTTCTTTT TGAAATGTTT
74701 CTTTCAAGTT CAAATTTTTT TCTTGTGTTT AGAGATCACA GTTAAGCAGC
74751 GTAGGTGGG AACTCAGAAC TACAATTGGA AAGCACTATT CTATTCAGGA
74801 CAGGATGTGG GAAGTGCTTT GATGGAATAT ATAAATTTGA TCTTAAGTAA
74851 TCAAGACATA GGTAGCAAAA AAATGGTCTA GGGCATCCCA ACAAATATAT
74901 GAATGCATTT GGGAAAATGT CATCATTTTG CTGGTGCTTA TGAAACTTGC
74951 TGTTGGGAGA TACCTATGCC TGTATGTAGG TATAGTTTAT AAGAAACCAT

FIGURE 3Y

75001 TTTAGAAAGA AAAGATAGAT TCCATTAATA CAGTTTTTAT CAGAGTACTT
75051 GGATTTTGGT TAATTCCTAC ATATTTTTTC TTAAACTTT TCTCAGTATT
75101 TTTATTGTTT AGAGAAATAA AACAAGATAA TCATTTTAAA TCATAGCACT
75151 TACGTTTTCT CTTGTTTTAT AAGGAGCAAG GATGCTCTAT AGAAAATATA
75201 ATGTAAGAAT AATAAAAGTT TTTGGTTTTT ACATAGGTAA AGCAACAGTG
75251 TGATTGGATT ATGGTGTTTG ATTCTATTCC ATTTTCAGCA AGAAAGCGAT
75301 GTTAACCACA AAGGAATAA GAAACATTTA AGATAGGCTG TGTGATTATG
75351 ATCTTTCAGA TCTTTGGCTC CTAATATCTG TTCCTTTATA TTCTATCACA
75401 CTCTTCTAAC TTTGGTAATC CTTGACAAAA GTGTGCACTT TATAACAAT
75451 CCTAAATCGA ATTGGTCTAT AGCTTAGAAT GGCTTTTTAA AGAATAATTG
75501 ATTCTGAGTA ATGTGGTCTG ATGAACAGTT TGATGACTTC AGTTTCTACT
75551 GAAAGAGAAG TTTCAAGTAT ACTGGCAACT ATATTCTGTT TTTTTCCTCC
75601 TGCAAAATAA GTTTAAATTT GGTTTGGGGG AAGGTTTGCC TTTATTTTTG
75651 CTTAATAAGG AGGCATTAGA AAGGGGCAGA GGAGGCTTGA CTGGTGTGTG
75701 CATTCTCTCC CTAGGTCGTC TTCAAAGGTG AAAAAAGGCG TGGTCACACT
75751 GGGGAGATTG GATTAGATGA TGTGAGCTTG AAAAAAGGCC ACTGCTCTGA
75801 AGAACGCTAA CAACTCCAGA ACTAACAATG AACTCCTATG TTGCTCTATC
75851 CTCTTTTTCC AATTCTCATC TTCTCTCCTC TTCTCCCTTT TATCAGGCCT
75901 AGGAGAAGG TGGGTCAGTG GGTGAGAAGG AAGTCTATTT GGTGACCCAG
75951 GTTTTTCTGG CCTGCTTTTG TGCAATCCCA ATGAACAGTG ATACCCTCCT
76001 TGAAATACAG GGGCATCGCA GACACATCAA AGCCATCTGT GGGTGTGACC
76051 TTCCATCCTG TGTCTCTTTC AGGAAGGCAT TCAGCATGCG TGAGCCATAC
76101 CATCCTCCAT CCTGATTACA AGGTGCTCCT TGAGCAAAT TATGAGAGTG
76151 ACTTACGGGA GCAGTTTTTA AAAGAAATCT TTGCAGATGG CTATGATGTT
76201 ATGTGTTTCG TGTGTACCA TGAGTAGTAT TGACTTCCCT TGAGATATGA
76251 TGTACAATGT GCTTGTGAAA TTGACTTACC CTCTTCACTT AAGTTAGTTC
76301 TGGCCTGACC TGAACCTCTGA CTTTTACTGC CATTCACTTT ATAAAATAAG
76351 GGTGTGTAA ACATCAAGAT ACATTTATTT TTATCTGTTT TTTTTTCTCCT
76401 GTTAAAGACA ATTATGTAGA GTGGGCACGT AATCCCTCCT TAGTAGTATT
76451 GTGTTTTGTG TAAATGTGCT ATTGATATTA AGTATTTACA TGTTCCAAAT
76501 ATTTACAGAC TCTAGTTGCA AGGTAAAGGG CAGCTTGTGA TCTCAAAAAA
76551 ATACATGGTG AAATGTCATC CAGTTCCATG ACCTTATATT GGCAGCAGTA
76601 GGAAATTGGC AGAAGTGTG GGTGTGGTA ACGGAGTGAT GAATTTTTTT
76651 TTAATGGCCT TGAGTTTGAT CTCTGCAAAG GATAGGAAAC CTTTAGGAAG
76701 ACAAGAAACT GCAGTTAATT TAGAACTGTC ACTGTTTCAA GTTACACTTT
76751 AAAACCACAG CTTTTACCAT CATACATGG CTCTGGTAAT ATGTAGGAAG
76801 CTTTATAAAA GTTTTGGTTG ATTCAGAAAA AGGATCCTGT TGCAGAGTGA
76851 GAGGAAGCAT AGGGGAAAC TCCATTGGAA CAGATTTTCA CACAACGTTT
76901 TAAATTGATA TAAATTTAGG CAGTTGTAGT TCATAACTTA TGTGCTCAT
76951 GTTGTGCTGT GTCAGGATGG GATAGGAAGC AAGTCCCATG CTTAGAGGCA
77001 TGGGATGTGT TGAACGGGA TTTACACACA CTGGAGGAGC AGGGCAAGTT
77051 GGAATCTTAA GATCCATGAA CCCCCAACTG TATTTCTCTC CTGCATATTT
77101 TACCAATATA TTAATAAACA ATGTAACCTT TAAAAGGCAT CATTCTGAG
77151 GTTTGTCTTA ATTTCTGATT AAGTAATCAG AATATTTTCT GCTGTTTTTG
77201 CCAGGAATCA CAAAGATGAT TAAAGGGTTG GAAAAAAGA TCTATGATGG
77251 AAAATTAAAG GAACTGGGAT TATTGAGCCT GGAGAAGAGA AGACTGAGGG
77301 GCAAACCATT GATGGTTTTT AAGTATATGA AGGGTTGGCA CAGAGAGGGT
77351 GGCAGCCAGC TGTCTCTCAT ATGCACTAAG AATAGAACAA GAGGAACTG
77401 GCTTAGACTA GAGTATAAGG GAGCATTTCT TGGCAGGGGC CATTGTTAGA
77451 ATACTTCATA AAAAAAGAAG TGTGAAAATC TCAGTATCTC TCTCTCTTTC
77501 TAAAAAATTA GATAAAATTT TGTCTATTTA AGATGGTTAA AGATGTTCTT
77551 ACCCAAGGAA AAGTAACAAA TTATAGAATT TCCCAAAAGA TGTTTTGATC
77601 CTACTAGTAG TATGCAGTGA AAATCTTTAG AACTAAATAA TTTGGACAAG
77651 GCTTAATTTA GGCATTTCCC TCTTGACCTC CTAATGGAGA GGGATTGAAA
77701 GGGGAAGAGC CCACCAAATG CTGAGCTCAC TGAAATATCT CTCCCTTATG
77751 GCAATCCTAG CAGTATTAAG GAAAAAGGA AACTATTTAT TCCAAATGAG
77801 AGTATGATGG ACAGATATTT TAGTATCTCA GTAATGTCCT AGTGTGGCGG
77851 TGGTTTTCAA TGTTTCTTCA TGCTAAAGGT ATAAGCCTTT CATTTGTTCA
77901 ATGGATGATG TTTTCTGATT TTTTTTTTTT AAGAGATCCT TCAAGGAACA
77951 CAGTTCAGAG AGATTTTCAT CGGGTGCATT CTCTCTGCTT CGTGTGTGAC

FIGURE 3Z

```

78001 AAGTTATCTT GGCTGCTGAG AAAGAGTGCC CTGCCCCACA CCGGCAGACC
78051 TTTCCTTCAC CTCATCAGTA TGATTCAGTT TCTCTTATCA ATTGGACTCT
78101 CCCAGGTTCC ACAGAACAGT AATATTTTTT GAACAATAGG TACAATAGAA
78151 GGTCTTCTGT CATTTAACCT GGTAAAGGCA GGGCTGGAGG GGGAAAATAA
78201 ATCATTAAAGC CTTTGAGTAA CGGCAGAATA TATGGCTGTA GATCCATTTT
78251 TAATGGTTCA TTTCCTTTAT GGTCATATAA CTGCACAGCT GAAGATGAAA
78301 GGGGAAAATA AATGAAAATT TTACTTTTCG ATGCCAATGA TACATTGCAC
78351 TAAACTGATG GAAGAAGTTA TCCAAAGTAC TGTATAACAT CTTGTTTATT
78401 ATTTAATGTT TTCTAAAATA AAAAATGTTA GTGGTTTTCC AAATGGCCTA
78451 ATAAAAACAA TTATTTGTAA ATAAAAACAC TGTTAGTAAT ACCAGTTGTC
78501 TATTCTTGTT TTTTGAGTTT TGTTTTTTTT TGACTTGGAA AAAAGCATTG
78551 AGGTAGTTAA ATGATGTTTC ACAAAGTCA TAGTAGAATC CCTTTTACTG
78601 PTTGGATGGT GGAACAAGG ATGTGCCTG CAGTATTATA CTTTCTAGGT
78651 TATAAACAT GAGACACTTT ATTTTTTTTA TCAGCATGAA CAGGGAAAGA
78701 GATCAGAAGA TCACTATAAC CCATGCCATG CCTTAGTAAA TTGCTTTAGT
78751 TATGTTTTAT TATCATTTCA TTGTAAACAT TTGCT (SEQ ID NO:3)

```

FEATURES:

Genewise results:

```

Start: 3000
Exon: 3000-3070
Exon: 5248-5348
Exon: 19533-19583
Exon: 34426-34518
Exon: 39701-39790
Exon: 44102-44221
Exon: 45394-45513
Exon: 47168-47302
Exon: 47607-47729
Exon: 49126-49127
Exon: 49132-49522
Exon: 65277-65363
Exon: 65834-65932
Exon: 74008-74264
Exon: 75715-75807
Stop: 75808

```

CHROMOSOME MAP POSITION:

chromosome 4

ALLELIC VARIANTS (SNPs):

DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
3365	C	T	Intron			
3411	G	A	Intron			
5243	G	A	Intron			
6532	T	G	Intron			
9208	-	G	Intron			
15833	T	C	Intron			
27795	G	T	Intron			
31341	A	G	Intron			
33370	G	A	Intron			
37788	G	A	Intron			
41465	C	T	Intron			
41466	A	G	Intron			
41653	T	G	Intron			
47666	G	A	Exon, coding	281	V	I
52613	G	A	Intron			

FIGURE 3AA

52645	G	A	Intron
59197	-	A	Intron
63508	T	C	Intron
75153	C	G	Intron
76017	C	G	Intron
76033	C	T	Intron
77194	A	G	Intron
77508	T	A	Intron
77557	G	A	Intron

Context:

DNA

Position

3365	CGGGAGGTGAGCTGGGCCCCGGGGCGCCCTCTCCTCCTTCCCGCGCTAATTTCACTCA CTGTCTTGGGTCACTTTTCCCCGCGGGTTTCGTGGTCAGAGAGGCGTCTCCTCCATCCA GAAGTTGGGCCACCGCACAGCGTGGCGCGAGGAGAGCGGTCCAGCGGCTCCGAGTGCCCCG CCCCAGGCGGAGAGGGCGGCCCTTGCGAGTCTGGGACCCATCCGCGGCCCCCGAGGG CGACTCGCCCCGCTCGGGAATTAGGACTGAGGGAGAGGAGCCGCTGGAGCCTGGGATCT [C,T] GGCTCTGAGGGCGCGGTTTAGCCACCTACGCCGAGGTGACGCGCGAAACATCCCTTACCC GGGAAACTCCCGCGCCTGAAGTAGACGGCTCTTCACTGGGGAAGCTTCCAGGCCCGGGG GGGAGGCCCGGGCTCTGCTCAGGGCTCTCGGGGCCGCTCACACAGAGAGTGGGTGCGAGT CAGCGACTGGGTACGGGGGAGATTTGTGGGCCTCTCCATTTGGTTTCTTGAGGGAAGG AGACTCAAAATGAGGACCGGAGGGTGGGCGCTCCGTGAATGTGAGCATGAGTGTGTGGAT
3411	TAATTTCACTCACTGTCTTGGGTCACTTTTCCCCGCGGGTTTCGTGGTCAGAGAGGC GTCTCCTCCATCCAGAAGTTGGGCCACCGCACAGCGTGGCGCGAGGAGAGCGGTCCAGCG GCTCCGAGTGCCCGCCGAGGCGGAGAGGGCGCGCCCTTGCGAGTCTGGGACCCATCCG CGGCCCCCGAGGGCGACTCGCCCCGGCTCGGGAATTAGGACTGAGGGAGAGGAGCCGCT GGAGCCTGGGATCTCGGCTCTGAGGGCGCGGTTTAGCCACCTACGCCGAGGTGACGCGCG [G,A] AACATCCCTTACCCGGGAAACTCCCGCGCCTGAAGTAGACGGCTCTTCACTGGGGAAGCT TCCAGGCCCGGGGGAGGCCCGGGCTCTGCTCAGGGCTCTCGGGGCCGCTCACACAGA GAGTGGGTGCGAGTCAGCGACTGGGCTACGGGGGAGATTTGTGGGCCTCTCCATTTGGTT TTCTTGAGGGAAGGAGACTCAAAATGAGGACCGGAGGGTGGGCGCTCCGTGAATGTGAGC ATGAGTGTGTGGATGTGTGTGTGAGAGCGCACACTGCGCCGCTCCTCAGACTCGGGCG
5243	GGCAGCCTTGAGAAATTGTTGGTGTTAGCTGGTGCTATTTCTACTTAAGAAAGCATCTTTC CCTCCCGTTTTTCTCCTACCCTTGTGTGTTGGGGGATGGGGTATTTAACTGTCTGAAAT TTGACATTAGATCATAGCAGAACTATTTCTGTGAAGGTGTTTAACTTAAATTTCTGCA GGCATAAAGAGTTTGTAGAAAATTGTTGGGTGTGGTTGACATTTTCTGAAGTAATAA TACAGAGGAAATTACCTTTCTCTCAAAAGTGTCCTTATTTATTTTGAATCTTTTTTTTT [G,A] GTAGGTGGCCAGGCAAATAGTGTGTCATCGATTGGCCTATGTCGTTATGGTGGGAGGATTG ACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTGAGCGTGAGTATCAAGCCT GGGACTTCAGTTCCCTGGGAGGTGTGGCTTTCACCTTGTTTCATGGCTTCACCCACAT ATCAGAGGGTTCACTACTGAGCAAGGCTTGGCCTTGCAGGTCTGACTTGGGGATTTCAG GTACAGTCCAGACTCCTTATCTGCTTCTTTTCAGCTTTAGCCACCTGTATTACGGCCCA
6532	GGGCGGGGTAGGGGGCACTTCTGTAAGATGATGGGAACTAAGTTGGGTTCTACATTGG GATATATATTTTATTGCTAATGAGGAGGAGGCTTAGAGGAAGAGAGAAGGGCAGTTACGA AGGCTAGAGCTGGCAATGGAGAAGCCTGCCTTAGAGATGGGTTGCTAGTGTGAGGAGTCA GGCAAATTTAAGTTCAAGAAAGTTAGGAGTTCCTCTGCTATTTTAAATTTTGAGGATGC TTGCAATGTCTTCTTAATTTGTGAAAGAGGGACAGTGACAGTCACAGATTGACTCTAA [T,G] TGCACATAAAGACCACAATCTCTGGTTGGGAATAGAAAGGTAAAGGAAATGAATGTTTGC CTACCTGGTATGGAATTTGAGAACCAACAGATTCTAATAACCAAAATGTGAAGAAAGGAC CCTTCTGTTGGCCCAACACACCTACACATAACCCTCCTGAGTGAAAAATGAGTAGTTCTA TACCTGCAGTCTCCAGCTGTGCAATACTTCTGATACTACAGAAGACTAAATCCACCAG

FIGURE 3BB

GCACCATTCTTCTTTTTTGATCATCTTCCCTTAAACAATATTGAATAGACTAACCAGTG

9208 TCTGAGTTCTTTGGCTGTCATTGAGCAACTTGTGGTTTCTGAAGGAAATTATGTGAATTA
GGATGGTTTTGTATCATTTATCCTTAAGAACAGGGGAAAATTGAGATGTTTTCTTATGTTT
CTGCTGGAGATTTTGGAAAGATGTGAAACCTACACCTACAGATTGACCTTGCTTAGT
CTCTGAACCTCCTGCTGCCTCTTCCACGTAAAGTGAAAATTTTGGATTCTTATCGGCTTC
AGATAAACTTACAGTTAGTGAAACATAGGGACTGAGATATAGTAATTCATTCTGAAGCT
[-,G]
TTTTGGAGTGGTCAAATAATTTTAGTTGGATAATATATATTACTGGCTAATGATTGTGGA
TATTGGAAGTGATGAAAAAATTATTGAATTATTTCTTCTGCATTTCAAATGAAAAGGCT
ATTAGTTTGAGCAGAGAATTTTGATTTAGTAAACAAAATATTTAAATTCATGTTTCATT
TCTTTCTCCTATCTGGGTTTCTGATCTCAGTCTTATAAATGGAACATGATTTATTTTTGC
TCCCTAAACTGGTTATTAACCTTCTGTCCATAATCACAAACTATATAGATTATATATTT

15833 CATACATTAGCTTTGGTAGATACTAGATGCCAGGGGAGTGCAAATTTGAGAATAAGGATT
TGGCATGGGTATTTGTCATGTTGAGAGTCAAGATTTAACTAAATTTTAAAACCAAACACT
TCATTTGATCAATATCTTTTTAACCATTCTGTAGATTAATAATATAAATTTCTCCAGAGCT
GACATTACTTTGCTATAACATCATCAGATCACAAGATTAGGGTTGCCTTTTGTAGATGTT
ATTCACTTACTCTGAAATGTTAGGAAATATGTCATCACAGTTTAAATTTGTAGTAATATA
[T,C]
ACAAAAGGAAAACACAACCTAGGAATTTTGGATTTATGCTTACTTTGCCAAAACCATGTT
GATTTTCAAAAACCTTTAGCC

27795 ATGAAAACAGAAAGAGAAAGTTAAAATTTCAAGTAAACTGAAAACAACAATAGTGCTCAG
AGGATTTTTTTTTAAGTGAAAAGGATAGTGCTTAATTATGACAAATAAAGCTAACTTG
AGATGCACATATACAGATGCAGCCACTTATTTTGGCGGGGACACTTCAGGAATTAATAAT
TTAAATAGCGAGAGATCAAATAGATACTTGGTAAATGTGCTGAATTGGATGTTCCCGA
CACAAAATAAATAGATTATTGACAGCTCTTGGGAGACAACATTATAAAGACTAGACAC
[G,T]
TTATTTATTTTAACTCTATGTTCTAAATTAACCATGAGTAATTGACATTCGTATTTGACT
ATGGTTTGTGGTTAAGTTCTTAATTGCAATAATGTTAAATAAATGTGAAGCCCAAAGCA
AACAACAACAAAATTATAGCAATACTTCAACAGAGGTAATAATAATATGCTGCATCAAT
GGTTTCAGAATCCAGATCTACATAAAAACAAGCAACAGGGTAATGAAATTTATTTCTTTT
AAATATTTCTGGCAGAGCTACTTTAGTTTCTTAAAGTTATAGATTGTGGTCTTAACCTGCAA

31341 CCACTTTGCTTATATTCTGAAAGTCTTTTGTTCCTATTCCACTGCTCTTACCTCTAACA
CACTGCCCTCTAACACACCAACCTGCAGTTGTAGTTATTACACACCCTCCTTGGTTCTTTC
ATCTCTCTATCACAGCCCTTGTGTGGTTTAGCCAATATATTTTAGTTCCACAGCTAAAT
TTTCATACCCTCTATGACTCTCTAATCCCCTGCCACACTTGCTTACTATAATACATTATA
TATATAACAAATGTTTGATACGTATTTATTGAATTCCATTCAGAACTAATGCCAGCAAG
[A,G]
TAACTTTGTGCTATATAGGAGAATATCTTTTTGTGCAACAGTTTCAAAGGGTTTTCTTT
TTCTAAGAAGAAAGAAATTGATTGTATCAACTTTATGAGTATCCTACCGCATTTAATAGC
CATTGGCTAATCTAAGGGTTCTGGTTACTTCACTGAATAGCCTATCAGATGGAAGTGCA
AACAACAGTTTGTGTTTGAATAGGACTCCCTAAACATGGAAGAAACATTAACAGTGTGG
CCTGTTGGAATGTGTGCATTTGATGTGCTCAAGATTAGGGCACTCTGCTTGAGAACAAT

33370 CCACTCACCTGATTGTTTCTCACTCTTCTCCACCATTACCCATCTCTTTAGCTTAATG
AGTCCCCCTGTGTATCTCCCACTAAACAGCGGCTTACTTGCTGTGAAATATTCTTCTCT
TGGGTAGTCTGCTCCCTTCTCTGTCTACTCATGCTTCAAGATTCAACATAAGCCTCCTCT
ATGAGGCTTTCTGCACGTATGTATATGGATTTGCTTGTGTAATGATTTCTTACAGATTT
CATATTGCTGATAAATAAATATTGTTTTGAATAAGAAACGTGGTTTTGTATTTTTATCTC
[G,A]
ATTGTAGACTCCTTGAGACCAGTACCATGCTATACAATTATTTTTCATCTATTATAGTGT
CTGGCATAGGGACATGCACATATTTGGTACAGAA

37788 GCCCTGTCAGGATCAGGGAAGCTGCCATGTGTGGCCTGGTTGAGTGTGAGAGCTGCCTAG
AGACTTCCATAAAAGTTGTTAGAGAAAATTGTTGATAGTGCCTACATAGCAGATTAACCT
AAACTGTTTTTCATTAGGCCACACATTTTAAATAAGTAGAAAATATGCTTCACAGATAA
GGGAAATCAAACAGGCTCCTTTTTTCTGGAGGAGAGAAATGTCAAAAAGAATTAAATTTG

FIGURE 3CC

AAATAACTTTACAGAACTGGAAATTAGCTTTTGATTAAAAGTAGCTTTTGGTATATGACA
[G, A]
GTATTCACTGAGAATTTTGTAGCGAGTTATATACTTTAAGAAATAACCCCCAGAACTTG
CATCATGGTGTAACAGCTTGAATAACAAGTGCTTAACCAGTGCCTTTAGAGCTGCCTG
GGAAACAGCCAGAATACCAGGGCAAGCTGCATTTTGGAACTGGTTAATTTAGTAGCCTT
GCCACAGGCTTAGTGTGATCTGCTTTTGGTGGCTTGATCTCCCCACTAAGTCATTTTCT
GGATTGTACACCTAGAAGTGTAGGAAATTACAGGCTTGGGCTGATCATTAACTACT

41465 TTTATTTCTAGGTTCTTTTATCAACTTTTAAGATTTATTTGAGACAGTATGATCAATGAC
TTCATTTTGTCTGCTTATTATTGTAGGAGTATTTACTATAATTTGGAGTAATTTATTT
TTGAATTTATTGCTTAATTGAATGATCTCCAATAGATTGTGATAATGAACACAGCATTTA
TAGAAAGCAGCAGATATTAACCTTACTTAATATGGCACTAGGTCAATGAGAAAAGAAGGTA
ACATAATTGAAGACAAGAACTCTTAAGAAAAGTGGGACAAAAAGGCTTCTCACCAGGA
[C, T]
ACCAGATGCATTTAATCTTTTGAAGCTCTGTACTTTAGGAAAAGTCTGATATTTGGCAAA
TTTTGATAAACATGGATGACTATGGAATCCTATTTTATAGTATCTGAAGTGGCTTTTATA
AGGGTCATTGTGAAGTTTTGTAGGAGACACCTGCCTGTGGCAGATGGGACAATGATGGCAG
TCACTACTGATATTAACACCAGTCAGCTGTGAGGAATATCATCCAGACCATCAGCAGCT
GGTAGAGTACAGCTTCTCAATTGCTTTCCATGTTTTGGTACTTATATGCCCCGTTAATA

41466 TTATTTCTAGGTTCTTTTATCAACTTTTAAGATTTATTTGAGACAGTATGATCAATGACT
TCATTTTGTTCGCTTATTATTGTAGGAGTATTTACTATAATTTGGAGTAATTTATTTT
TGAATTTATTGCTTAATTGAATGATCTCCAATAGATTGTGATAATGAACACAGCATTAT
AGAAAGCAGCAGATATTAACCTTACTTAATATGGCACTAGGTCAATGAGAAAAGAAGGTA
CATAATTGAAGACAAGAACTCTTAAGAAAAGTGGGACAAAAAGGCTTCTCACCAGGAC
[A, G]
CCAGATGCATTTAATCTTTTGAAGCTCTGTACTTTAGGAAAAGTCTGATATTTGGCAAA
TTTGATAAACATGGATGACTATGGAATCCTATTTTATAGTATCTGAAGTGGCTTTTATA
GGGTCAATTGTGAAGTTTTGTAGGAGACACCTGCCTGTGGCAGATGGGACAATGATGGCAGT
CACTAGTGATATTAACACCAGTCAGCTGTGAGGAATATCATCCAGACCATCAGCAGCTG
GTAGAGTACAGCTTCTCAATTGCTTTCCATGTTTTGGTACTTATATGCCCCGTTAATA

41653 AGCACATATTAACCTTACTTAATATGGCACTAGGTCAATGAGAAAAGAAGGTAACATAATT
GAAGACAAGAACTCTTAAGAAAAGTGGGACAAAAAGGCTTCTCACCAGGACACCAGAT
GCATTTAATCTTTTGAAGCTCTGTACTTTAGGAAAAGTCTGATATTTGGCAAAATTTGAT
AAACATGGATGACTATGGAATCCTATTTTATAGTATCTGAAGTGGCTTTTATAAGGGTCA
TTGTGAAGTTTTGTAGGAGACACCTGCCTGTGGCAGATGGGACAATGATGGCAGTCACTAG
[T, G]
GATATTAACACCAGTCAGCTGTGAGGAATATCATCCAGACCATCAGCAGCTGGTAGAGT
ACAGCTTTCTCAATTGCTTTCCATGTTTTGGTACTTATATGCCCCGTTAATAACAGGTAA
AATAGCCAGTACATCATTTCCACATTTACCCATTGAATGTTGCATGTTTTCTTCTTTCA
CATATTCATACAGTCCAGATTTTGTGGACTCATGACAGCAGCATTGGCTTTTCTTTCC
TTTCAGTTTCATGATTCTTAACCCCAAAGTGCTTTTGCCATGGGAACGGAAGGATAAATT

47666 GTGGTGATGGCTGGAATGTCAGGGGAGGGGAGAGTACTGGCGTTAAGTTAAACCAACAG
ACATCCAGTTTAAACCACTGGTAGTTCTCAGTCTACATGTAGTTTATTTCTTCTGTTTATC
TGCCAATTTTATGTAGATCATCACATTGCCAAAAAAATCATTTTGAAGTGTATATAT
TTTTTATGTCATCATATTTATCTCCTAAATAAGTCTCTTCTTTTCTACTTTCTGATGCA
GACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAAC
[G, A]
TACGTGGGTCTTACAAGTGCAATGTAAGAAGGATACCAGGGTGATGGACTGACTTGTG
TGTGTGAGTAGCACTTGTCTCTCAGCTTTAAATTCTAGCAGGAAATACAGGATTACACAA
AGGCCATTGCTAGGGAATAAGGAATAAGATTATCAAGAAGTATAATTGTCATAATTG
GTTATATTTGCTTTTGAATTTCCACAAACAATAAAATCACTTGTCTCAGGTACTTGTAAAA
CTAAGGACTCAGTAATACACTATAATCTTAAGAGTATTTAATCTCTTCACTGAAATCTC

52613 AGGTAATGGGACCTCTATCACCTGTTGCTAAGGTCTAATAATGTGGATTTCATCCTTAATT
ACTCTTCATACATACAGTCAATATGTAAGGAAATCCTGTTAGCTGTACCTTCAAATTATA
TTTAAGTGTGACCTTTTCTCACCACCTCCAGTGCTACCACCTGGCCCAAGCCATCTCTC
CCCTGGAATACAGGCCCAAAATCCTTTATCCAAATAACTTATGGAATATAGCATTTTTTA

FIGURE 3DD

GATTTTAGAAAATCAGTAAGGTACACATACTATATAGTACACACTGAAGTAGTGAAACAT
 [G, A]
 CTGATTTTCCTCTAGTGCTTTTACTGTGAACGTATCAATGTTAAGGAAAGGAAAATGATA
 TTAACCTCAGAGATGATGTCTCACAGCATATATTTACTAGCTTGCACAAATTTTTAATGT
 TAGCAAGATTTAAGGCAAATTTTTACTTATATTTAATTGGATCCTATGATGATTATTA
 AAGAAAAAAGTAGTTATCTCTAGAAAGTATAAATTGAGCTCTTGGCAAATGTGAGCAAAA
 CCAGAAATCAGATTTTTTTAAAGTTACGTGTACATTTGTCTATAGAGTTATAATTAAAG

52645 GTCTAATAATGTGGATTTCATCCTTAATTACTCTTCATACATACAGTCAATATGTAAGGAA
 ATCCTGTAGCTGTACCTTCAAATTATATTTAAGTGTGACCTTTTCTCACCACCTCCAGT
 GCTACCACCCTGGCCCAAGCCATCTCTCCCTGGAATACAGGCCCAAAATCCTTTATCCA
 AATAACTTATGGAATATAGCATTTTTTAGATTTTAGAAAATCAGTAAGGTACACATACTA
 TATAGTACACACTGAAGTAGTGAAACATGCTGATTTTCCTCTAGTGCTTTTACTGTGAAC
 [G, A]
 TATCAATGTTAAGGAAAGGAAAATGATATTAACCTCAGAGATGATGTCTCACAGCATATAT
 TTACTAGCTTGCACAAATTTTTTAATGTTAGCAAGATTTAAGGCAAATTTTTTACTTATA
 TTTTAATTGGATCCTATGATGATTATTAAGAAAAAAGTAGTTATCTCTAGAAAGTATAA
 ATTGAGCTCTTGGCAAATGTGAGCAAAACCAGAAATCAGATTTTTTTAAAGTTACGTGTA
 CATTTGTCTATAGAGTTATAATTAAAGTATTGTGCTCATAGCCAGTCTGTGATCTGGGG

59197 CTACTATTGATGTTGAATTCATTTATATAGTAATGATAACATTTCTACTTAATTCATAA
 AAAGACAGCCTATGCTGTTTTCTGTCTGAGTTTATATGTTTCTCATGCTTTTTATTAT
 GGTTCATTACAATTTAATGTTATTTTTAACTAACTAGATCCTTTTGAAACAAATTGGTT
 TGCAAGTGTGAGCTGTTAGGTGCACAGAGAAAAATGAAATAGAACTTGCGATTTTATT
 CTAGGCTTGTTACCAAATATTAGAATACTGTGTTTTATTAGGTGTTTATAGTCTCATT
 [-, A]
 GACAGTTGTGATTTTAAATAGAGACCACATCATCTCAACTCTTTACTGTGAAAATAAT
 GACAATAGTCTTTTCAGAGATGAATCTGTCTAGATGGGAAATTTACATGATTGATCTGAT
 GAGNN
 AAAATAATCTGACAAGTAGTTTCCCCAGAAAATCTGATTTAGTAAATGTACCAAAAGGAT
 TTAGAAATCTACATCATAAATAAACATTCTATGTTATTTTAGTTTCAGACCCTATTTTAAT

63508 ACAGGCTCCCGGTGCCAGGAAACCTTTACATCAACACTCGATTTGCCATTTGATAGTCTT
 TCATCTGGGAGGAAAAAAGACGAGGGGAGCTTGAAAACTGTCATAATGTCCCTGG
 AATATGGTACTTTTAAGAGTTGAGCCTATTCCATTTTGGAGATGATTTATATAAGTTACA
 ACAAAAGAAGGGGACAAAACATGATTGTTCTATGGAGTTTATAACTTTCTGTACAA
 GAAAGCACGCTTGCTACAATTTTGTAATATTCTAGTAAATAAAGAGGCACTCCCGGT
 [T, C]
 TCAGAGCACCAAATAAGGAAAGTGAATTTGGATGTCATTGCTGTGAGTCACTGGGCTAT
 AAAAGAGAGAGTGGGGTTGCCTCATCCCTGGGTATCCACAGTCAGCTGTGTCCCTAGAG
 CTTCTTTTCTTTTCATTGCTGCCCAGCTGGGTATATTGCAAGTATGGATTATAAGAGGGGA
 AGGGACTTCACTGTTTAAACGTTTGAAACAAAAGGAAAAAATCAGAAAGTAGTAAGCTA
 AAAACAATCTGTGCAAACGTTCTGGGATTATTACTTAATTTTAAGAATTTTGTCTAAAA

75153 AAGACATAGGTAGCAAAAAAATGGTCTAGGGCATCCCAACAAAATTATGAATGCATTTGG
 GAAAATGTCATCATTTTGTGGTGCTTATGAACTTGCTGTTGGGAGATACCTATGCCTG
 TATGTAGGTATAGTTTATAAGAAACCATTTTAGAAAGAAAAGATAGATTCCATTAATACA
 GTTTTTATCAGAGTACTTGGATTTTGTGTTAATTCTTACATATTTTTTCTTAAACTTTTC
 TCAGTATTTTTATTGTTTAGAGAAATAAACAAGATAATCATTTTAAATCATAGCACTTA
 [C, G]
 GTTTTCTCTTGTTTTATAAGGAGCAAGGATGCTCTATAGAAAAATAATGTAAGAATAAT
 AAAAGTTTTTGGTTTTTACATAGGTAAAGCAACAGTGTGATTGGATTATGGTGTGTTGATT
 CTATTCCATTTTCAGCAAGAAAGCGATGTTAACCACAAAGGAACTAAGAAACATTTAAGA
 TAGGCTGTGTGATTATGATCTTTTCAGATCTTTGGCTCCTAATATCTGTTCCCTTTATATTC
 TATCACACTCTTCTAACTTTGGTAATCCTTGACAAAAGTGTGCACTTTATAAACAATCCT

76017 CGTCTTCAAAGGTGAAAAAAGGCGTGGTCACTGAGGAGATTGGATTAGATGATGTGAG
 CTTGAAAAAAGGCCACTGCTCTGAAGAAGCTAACAACCTCAGAACTAACAATGAACCTCC
 TATGTTGCTCTATCCTCTTTTTCCAATTCTCATCTTCTCTCCTCTCTCCCTTTTATCAG
 GCCTAGGAGAAGAGTGGGTGAGTGGGTGAGGAGTCTATTGGTGACCCAGGTTTTT

FIGURE 3EE

CTGGCCTGCTTTTGTGCAATCCCAATGAACAGTGATACCCTCCTTGAAATACAGGGGCAT
[C, G]

GCAGACACATCAAAGCCATCTGTGGGTGTTGCCTTCCATCCTGTGTCTCTTTCAGGAAGG
CATTACAGCATGCGTGAGCCATACCATCCTCCATCCTGATTACAAGGTGCTCCTTGTAGCA
AATTATGAGAGTGAGTTACGGGAGCAGTTTTTAAAAGAAATCTTTCGAGATGGCTATGAT
GTTATGTGTTTCGGTGTGTACCATGAGTAGTATTGACTTCCCTTGAGATATGATGTACAA
TGTGCTTGTGAAATTGACTTACCCTCTTCACTTAAGTTAGTTCTGGCCTGACCTGAACTC

76033 AAAAGGCGTGGTCACACTGGGGAGATTGGATTAGATGATGTGAGCTTGAAAAAGGCCAC
TGCTCTGAAGAACGCTAACAACCTCCAGAACTAACAATGAACTCCTATGTTGCTCTATCCT
CTTTTCCAATTCTCATCTTCTCCTCTTCTCCCTTTTATCAGGCCTAGGAGAAGAGTG
GGTCAGTGGGTGAGAAGGAAGTCTATTTGGTGACCCAGGTTTTTCTGGCCTGCTTTTGTG
CAATCCCAATGAACAGTGATACCCTCCTTGAAATACAGGGGCATCGCAGACACATCAAAG
[C, T]

CATCTGTGGGTGTTGCCTTCCATCCTGTGTCTCTTTCAGGAAGGCATTACGCATGCGTGA
GCCATACCATCCTCCATCCTGATTACAAGGTGCTCCTTGTAGCAAATTATGAGAGTGAGT
TACGGGAGCAGTTTTTAAAAGAAATCTTTCGAGATGGCTATGATGTTATGTGTTTCGGTGT
TGTACCATGAGTAGTATTGACTTCCCTTGAGATATGATGTACAATGTGCTTGTGAAATTG
ACTTACCCTCTTCACTTAAGTTAGTTCTGGCCTGACCTGAACTCTGACTTTTACTGCCAT

77194 AACGTTTTAAATTGATATAAGTTTAGGCAGTTGTAGTTCATAAATTATGTTGCTCATGTT
GTGCTGTGTGAGGATGGGATAGGAAGCAAGTCCCATGCTTAGAGGCATGGGATGTGTGG
AACGGGATTTACACACACTGGAGGAGCAGGGCAAGTTGGAATTCTAAGATCCATGAACCC
CCAACGTATTTCTCCCTGCATATTTACCAATATATTAATAAACAATGTAACTTTTAA
AAGGCATCATTCCTGAGGTTTGTCTTAATTTCTGATTAAAGTAATCAGAATATTTTCTGCT
[A, G]

TTTTTGGCAGGAATCACAAGATGATTAAAGGGTTGAAAAAAGATCTATGATGGAAAA
TTAAAGGAACTGGGATTATTGAGCCTGGAGAAGAGAAGACTGAGGGGCAAACCATTGATG
GTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTGTTCTCCATATGC
ACTAAGAATAGAACAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTTCTTGGC
AGGGGCCATTGTTAGAATACTTCATAAAAAAAGAAGTGTGAAAATCTCAGTATCTCTCTC

77508 TCACAAAGATGATTAAAGGGTTGAAAAAAGATCTATGATGGAAAATTAAAGGAACTGG
GATTATTGAGCCTGGAGAAGAGAAGACTGAGGGGCAAACCATTGATGGTTTTCAAGTATA
TGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTGTTCTCCATATGCACTAAGAATAGAA
CAAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTTCTTGGCAGGGGCCATTGTT
AGAATACTTCATAAAAAAAGAAGTGTGAAAATCTCAGTATCTCTCTCTCTTTCTAAAAAA
[T, A]

TAGATAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACA
AATTATAGAATTTCCCAAAAGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTT
AGAACTAAATAATTTGGACAAGGCTTAATTTAGGCATTTCCCTCTTGACCTCCTAATGGA
GAGGGATTGAAAGGGGAAGAGCCACCAAATGCTGAGCTCACTGAAATATCTCTCCCTTA
TGGCAATCCTAGCAGTATTAAAGAAAAAGGAACTATTTATTCCAAATGAGAGTATGAT

77557 AAAGGAACTGGGATTATTGAGCCTGGAGAAGAGAAGACTGAGGGGCAAACCATTGATGGT
TTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTGTTCTCCATATGCAC
TAAGAATAGAACAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTTCTTGGCAG
GGGCCATTGTTAGAATACTTCATAAAAAAAGAAGTGTGAAAATCTCAGTATCTCTCTCTC
TTTCTAAAAAATTAGATAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAA
[G, A]

GAAAAGTAACAAATTATAGAATTTCCCAAAAGATGTTTTGATCCTACTAGTAGTATGCAG
TGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGGCATTTCCCTCTTGAC
CTCCTAATGGAGAGGGATTGAAAGGGGAAGAGCCACCAAATGCTGAGCTCACTGAAATA
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAAGAAAAAGGAACTATTTATTCCAAAT
GAGAGTATGATGGACAGATATTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTT

FIGURE 3FF